# KNOWING WHERE TO START: EIF2BETA AND RPS26/ES26 AS DETERMINANTS OF ACCURATE AND EFFICIENT START CODON SELECTION FOR TRANSLATION INITIATION

by
Laura Marler

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
July 2019

# Abstract

Translation initiation is the critical first step of protein synthesis in which a start codon is selected, determining the identity and quantity of proteins produced. In eukaryotes, the 40S small ribosomal subunit forms a 43S preinitiation complex (PIC) with eukaryotic initiation factors eIF1, eIF1A, eIF3, eIF5, and a ternary complex (TC) of GTP-bound eIF2 and methionyl initiator tRNA (Met-tRNA$_i$), which is recruited to the 5' end of mRNA, and scans for an AUG start codon in appropriate sequence context. Upon start codon recognition, release of eIF1 and completion of eIF2-GTP hydrolysis accompany structural changes in the 40S subunit that precipitate the cessation of scanning. The precise molecular mechanism of start codon and sequence context recognition is unknown.

Structural evidence indicates that eIF2β may stabilize Met-tRNA$_i$ and eIF1 in the open PIC. eIF2β-S202A/K214A, which disrupts eIF2β:Met-tRNA$_i$ interaction, decreases initiation accuracy and TC loading *in vivo* and a defect in TC on-rate and increased transition to the closed state *in vitro*, consistent with destabilization of the open complex. eIF2β-E189R alleviates a clash between eIF2β and Met-tRNA$_i$ and confers a hypoaccuracy phenotype *in vivo* and increased transition to the closed state *in vitro* without affecting TC loading. These mutants represent two mechanisms by which eIF2β prevents inappropriate transition to the closed state at non-AUG codons: by stabilizing Met-tRNA$_i$ binding in the open complex and by hindering rearrangement of Met-tRNA$_i$

to its closed state position. eIF2β -F217A/Q221A, designed to disrupt the eIF2β:eIF1 interface, produced a hypoaccuracy phenotype and a mild defect in TC loading, suggesting these residues aid eIF1 binding on the open complex.

Substitutions of well conserved residues in small subunit ribosomal protein Rps26/eS26 likely to contact the -10 to -12 nucleotides resulted in increased stringency of the requirement for an AUG start codon and appropriate sequence context, indicating that Rps26/eS26 contacts with mRNA in the exit channel may help to stabilize the closed state following AUG recognition. Our findings add to a growing network of molecular interactions that ensure the accuracy and efficiency of translation initiation, maintaining precise control of cellular protein levels.

Thesis advisor and primary reader: Dr. Alan G. Hinnebusch (Senior Investigator, Eunice K. Shriver National Institute of Child Health and Human Development, National Institutes of Health)

Thesis committee member and secondary reader: Dr. Kyle Cunningham (Professor, Department of Biology, Johns Hopkins University)

Thesis committee members: Dr. Tom Dever (Senior Investigator, Eunice K. Shriver National Institute of Child Health and Human Development, National Institutes of Health); Dr. Jon Dinman (Professor, Department of Biology, University of Maryland, College Park)

"He who chooses the beginning of the road chooses the place it leads to."
-Harry Emerson Fosdick

# Acknowledgements

This work would not have been possible without the support and assistance of my advisor, colleagues, and collaborators. I am deeply grateful to my mentor, Alan Hinnebusch, for his unwavering support throughout my work, his invaluable intellectual contributions, and his enthusiasm for these projects, as well as his patience, understanding, and encouragement. I would like to thank the members of the Hinnebusch lab, as well as many past members, for their guidance and help, and for being friends as well as colleagues. I would especially like to thank Fan Zhang, who was always available for a discussion and always ready to help me climb the next mountain (be it metaphorical or literal); Jyothsna Visweswaraiah, who spent untold hours patiently explaining her work to me with grace and humor, and still found time to explore a castle, a new restaurant, or a new recipe with me; Anil Thakur, who was a thoughtful and inquisitive coauthor, collaborator, and friend; and Quira Zeidan, who was the best benchmate that I could have asked for. I am also appreciative of the many contributions of my collaborators, Jose Llacer, Tanweer Hussein, and Venki Ramakrishnan.

I would like to acknowledge the members of my committee, Tom Dever, Kyle Cunningham, and Jon Dinman, for their invaluable contributions and their encouragement. I am grateful to Jon Lorsch and Nick Guydosh for their thoughtful comments and suggestions at lab meetings, and to the members of their labs, as well as the Dever lab members, for allowing me the opportunity to learn from their work. In particular, I would like to thank Jagpreet Nanda, who dedicated many hours to teaching me biochemical techniques. I am indebted to the staff of the NIH OITE and the Graduate

Partnerships Program for their financial and emotional support, and to Orna Cohen-Fix and Michael Lichten for the energy they poured into making the Johns Hopkins GPP excellent. I am grateful to John Pezzuto, whose training and encouragement were vital to my decision to pursue a graduate career. I would also like to thank Elaine Johnson, Bobbie Felix, Joan Miller, and the members of the CMDB office staff for outstanding administrative support.

Finally, I would like to thank my family and friends, without whom, the completion of this dissertation would not have been possible. I cannot begin to express my gratitude to my parents, Joseph and Barbara Marler, who awakened my interest in the natural world and demonstrated unwavering support throughout my education. I am incredibly grateful for the support and encouragement of my husband, Matthew Rice, who helped me through each stage of my PhD training and who, further, shouldered considerable responsibilities to create the time and space for me to write. I am thankful for my classmates at Johns Hopkins, and for the many friends who helped me through this process by commiserating, celebrating my successes, or providing a happy distraction. Thank you all.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| $^{14}C$ | **C**arbon-**14** |
| $^{35}S$ | **S**ulphur-**35** |
| 4E-BP | eIF**4E**-**B**inding **P**rotein |
| 5-FOA | **5**-**F**luoro**o**rotic **A**cid |
| 5MP | eIF**5** **M**imic **P**rotein |
| 5q- | Deletion of the long (**q**) arm of chromosome **5** |
| µL | Microliter |
| µm | Micromolar |
| Å | Angstrom |
| A | **A**denine (nucleotide); **A**lanine (protein residue) |
| $A_{600}$ | **A**bsorbance at **600** nm |
| AD | **A**lzheimer's **D**isease |
| Ala | **Ala**nine |
| Arg | **Arg**inine |
| A site | **A**minoacyl **site** |
| ASL | **A**nticodon **S**tem **L**oop |
| Asp | **Asp**artic Acid |
| ATP | **A**denosine **T**ri**p**hosphate |
| bp | **B**ase **P**air |
| bS | **B**acterial **S**mall Subunit Ribosomal Protein |
| C | **C**ysteine |
| CDS | **Co**d**d**ing **S**equence |
| CITE | **C**ap **I**ndependent **T**ranslation **E**lement |
| cp | **C**entral **P**rotuberance of the LSU |
| Cryo-EM | **Cryo**-**E**lectron **M**icroscopy |
| CTD | **C**-**T**erminal **D**omain |
| CTT | **C**-**T**erminal **T**ail |
| D | Aspartic Acid |
| d | **D**ays |
| DHRC | **D**irected **H**ydroxyl **R**adical **C**leavage |
| DNA | **D**eoxyribo**n**ucleic **A**cid |
| E | Glutamic acid |
| *E. coli* | ***Escherichia coli*** |
| EDTA | **E**thylene**d**iamine**t**etra**a**cetic Acid |
| EF | **E**longation **F**actor |
| eIF | **E**ukaryotic **I**nitiation **F**actor |
| el.uORF | **El**ongated **U**pstream **ORF** |
| EMSA | **E**lectrophoretic **M**obility **S**hift **A**ssay |
| ER | **E**ndoplasmic **R**eticulum |
| eS | **E**ukaryotic **S**mall Subunit Ribosomal Protein |
| ES | **E**xpansion **S**egment |

| | |
|---|---|
| E site | **E**xit **site** |
| F | Phenylalanine |
| Fl | **Fl**uorescent |
| fMet-tRNA$_i$ | **F**ormyl-**Met**hionyl **I**nitiator **tRNA** |
| FMRP | **F**ragile X **M**ental **R**etardation **P**rotein |
| FRET | **F**luorescence **R**esonance **E**nergy **T**ransfer |
| FSC | **F**ourier **S**hell **C**orrelation |
| G | **G**uanine |
| Gal | **Gal**actose |
| GAP | **G**TPase **A**ctivating **P**rotein |
| Gcd$^-$ | **G**eneral **C**ontrol **D**erepressed |
| GDF | **G**DI **D**issociation **F**actor |
| GDI | **G**DP **D**issociation **I**nhibitor |
| GDP | **G**uanosine **Di****p**hosphate |
| GDPNP | Nonhydrolyzable GTP analog |
| GEF | **G**uanine Nucleotide **E**xchange **F**actor |
| Glu | **Glu**tamic Acid |
| G-protein | **G**uanine Nucleotide Binding **Protein** |
| GTP | **G**uanosine **T**ri**p**hosphate |
| H | **H**istidine |
| h | **H**elix of 18S rRNA (e.g., h28); **H**ours (time) |
| hc | **H**igh **C**opy (plasmid) |
| HEAT domain | **H**untingtin, elongation factor 3 (**E**F3), protein phosphatase 2A (PP2**A**), and **T**OR1 |
| Hepes-KOH | (4-(2-**H**ydroxy**e**thyl)-1-**P**iperazine**e**thane**s**ulfonic Acid)-Potassium Hydroxide |
| His | Histidine |
| HRI | **H**eme-**R**egulated e**I**F2α Kinase |
| HTH | **H**elix-**T**urn-**H**elix Domain |
| I | **I**soleucine |
| IC | **I**nitiation **C**omplex |
| IF | **I**nitiation **F**actor |
| IRES | **I**nternal **R**ibosome **E**ntry **S**ite |
| IRP | **IR**E-binding **P**rotein |
| ISR | **I**ntegrated **S**tress **R**esponse |
| K | Lysine |
| kb | **K**ilo**b**ase |
| KCl | Potassium Chloride |
| $K_d$ | Equilibrium Dissociation Constant |
| kDa | **K**ilo**da**ltons |
| *K. lactis* | ***K****luyveromyces* ***lactis*** |
| $k_{obs}$ | Observed rate (pseudo-second order rate constant) |
| $k_{off}$ | Dissociation rate |
| $k_{on}$ | Association rate |
| kV | **K**ilo**v**olts |

| | |
|---|---|
| L | **L**eucine |
| L1 | **L1** stalk of the LSU |
| lc | **L**ow **C**opy (plasmid) |
| Leu | **Leu**cine |
| LSU | **L**arge **S**ub**u**nit |
| M | **M**ethionine |
| $m^6A$ | *N*$^{\mathbf{6}}$**-M**ethyl**a**denosine |
| $m^7G$ | **7-M**ethyl**g**uanylate Cap |
| Mat | **Mat**ing Type |
| MEHMO | **M**ental Retardation, **E**pilepsy, **H**ypogenitalism, **M**icrocephaly, **O**besity |
| Met | **Met**hionine |
| Met-$tRNA_i$ | **Met**hionyl **I**nitiator **tRNA** |
| MFC | **M**ulti-**F**actor **C**omplex |
| $MgCl_2$ | Magnesium Chloride |
| mM | **M**illi**m**olar |
| mRNA | **M**essenger **RNA** |
| mTOR | **M**ammalian **T**arget **o**f **R**apamycin |
| N | Asparagine |
| NaF | Sodium Fluoride |
| nm | **N**ano**m**olar |
| NMR | **N**uclear **M**agnetic **R**esonance |
| nt | **N**ucleo**t**ide |
| NTD | **N-T**erminal **D**omain |
| NTT | **N-T**erminal **T**ail |
| OB | **O**ligonucleotide/oligosaccharide **B**inding Fold |
| ORF | **O**pen **R**eading **F**rame |
| PABP | **P**oly(**A**)-**B**inding **P**rotein |
| PCI | **P**rotease/**C**op9/e**I**F3 domain |
| PCR | **P**olymerase **C**hain **R**eaction |
| PD | **P**arkinson's **D**isease |
| PDB | **P**rotein **D**ata**b**ank |
| PERK | **P**KR-like **E**ndoplasmic **R**eticulum **K**inase |
| Phe | **Phe**nylalanine |
| $P_i$ | **I**norganic **P**hosphate |
| PI3K | **P**hosphotidyl**i**nositol 3-**K**inase |
| PIC | **P**re**i**nitiation **C**omplex |
| $P_{IN}$ | Position of Met-$tRNA_i$ bound deeply in the P site in the closed PIC |
| PKR | **P**rotein **K**inase **R** |
| P/M | **P**olysome to **M**onosome Ratio |
| pm | **P**artial **M**ammalian (structure) |
| PMSF | **P**henyl**m**ethanyl**s**ulfonyl **F**luoride |
| $P_{OUT}$ | Position of Met-$tRNA_i$ when not fully engaged in the P site in the open PIC |

| | |
|---|---|
| P site | **P**eptidyl **site** |
| PTC | **P**eptidyl **T**ransferase **C**enter |
| py | **P**artial **Y**east (structure) |
| Q | Glutamine |
| R | Arginine |
| RACK1 | **R**eceptor for **A**ctivated **C** **K**inase **1** |
| RAN | **R**epeat **A**ssociated **N**on-AUG Translation |
| RAS Syndrome | **R**edundant **A**cronym **S**yndrome |
| RMSD | **R**oot-**M**ean-**S**quare **D**eviation |
| RNA | **R**ibo**n**ucleic **A**cid |
| Rpl | **R**ibosomal **P**rotein of the **L**arge Subunit |
| r-protein | **R**ibosomal **Protein** |
| Rps | **R**ibosomal **P**rotein of the **S**mall subunit |
| rRNA | **R**ibosomal **RNA** |
| S | **S**vedberg unit (measure), **S**erine (protein residue) |
| s | **S**econds |
| S1P | **S**phingosine-**1**-**p**hosphate |
| S6K | p70 **S6 K**inase |
| SC | **S**ynthetic **C**omplete (glucose medium) |
| sc | **S**ingle **C**opy (plasmid) |
| *S. cerevisiae* | ***S****accharomyces* ***cerevisiae*** |
| SD | **S**ynthetic **D**extrose (minimal glucose medium); **S**tandard **D**eviation (statistics) |
| SDS-PAGE | **S**odium **D**odecyl **S**ulfate – **P**oly**a**crylamide **G**el **E**lectrophoresis |
| SE | **S**canning **E**nhancer Element |
| SEM | **S**tandard **E**rror of the **M**ean |
| SI | **S**canning **I**nhibitor Element |
| Slg$^-$ | **Sl**ow **G**rowth |
| SSU | **S**mall **Sub**unit |
| Ssu$^-$ | **S**upressor of **Su**i- |
| St | L7/L12 **St**alk of the LSU |
| Sui$^-$ | **Su**pressor of **I**nitiation Codon |
| $t^6A$ | $N^6$-**t**hreonylcarbamoyl-**a**denosine |
| TC | **T**ernary **C**omplex |
| TCA | **T**ri**c**hloro**a**cetic Acid |
| TIR | **T**ranslation **I**nitiation **R**egion |
| TISU | **T**ranslation **I**nitiator of **S**hort 5'-**U**TR |
| Tris-HCl | **Tris**(hydroxymethyl)aminomethane-**H**ydro**c**h**l**oric Acid |
| tRNA | **T**ransfer **RNA** |
| $tRNA_e$ | **E**longator **tRNA** |
| $tRNA_i$ | **I**nitiator **tRNA** |
| Trp | **Tryp**tophan |
| U | **U**racil |
| uORF | **U**pstream **O**pen **R**eading **F**rame |
| Ura | **Ura**cil |

| | |
|---|---|
| uS | **U**niversal **S**mall Subunit Ribosomal Protein |
| UTR | **Unt**ranslated **R**egion |
| VR | **V**ariable **R**egion |
| WCE | **W**hole **C**ell **E**xtract |
| WT | **W**ild **T**ype |
| Y | Tryptophan |
| ZBD | **Z**inc **B**inding **D**omain |
| | |
| Gene Names: | |
| *ATF4* | **A**ctivating **T**ranscription **F**actor **4**, a mammalian homolog of *GCN4* |
| CDKN2A | **C**yclin **D**ependent **K**inase I**n**hibitor **2A**, Encoding p16 (mammals) |
| CYFIP1 | **Cy**toplasmic **F**MR1 **I**nteracting **P**rotein **1** (mammals) |
| Ded1/Ddx3 | **D**efines **E**ssential **D**omain 1 (yeast)/ **D**EA**D**-bo**x** Helicase 3 (mammals) |
| Dhx29 | **D**Ex**H**-bo**x** Helicase 29 (mammals) |
| *EIF2S3* | Encodes eIF2γ (mammals) |
| *FUN12* | Encodes eIF5B (yeast) |
| *GADD34* | **G**rowth **A**rrest and **D**NA **D**amage Inducible Protein **34** (mammals) |
| *GCD11* | Encodes eIF2γ (yeast) |
| *GCN2* | Kinase that phosphorylates eIF2α (yeast) |
| *GCN4* | Transcriptionally regulated activator of amino acid biosynthetic genes (yeast) |
| GRP78/BiP | **B**inding **I**mmunoglobulin **P**rotein (mammals) |
| *Hcr1* | Encodes eIF3j (yeast) |
| *HIS4* | Encodes multifunctional enzyme required for histidine biosynthesis (yeast) |
| *lacZ* | Encodes β-galactosidase (bacteria) |
| *LEU2* | Encodes β-isopropyl malate dehydrogenase, required for leucine biosynthesis (yeast) |
| *SUI1* | Encodes eIF1 (yeast) |
| *SUI2* | Encodes eIF2α (yeast) |
| *SUI3* | Encodes eIF2β (yeast) |
| *TRP1* | Encodes Phosphoribosylanthranilate isomerase, required for tryptophan biosynthesis (yeast) |
| *URA3* | Encodes Orotodine-5'-phosphate decarboxylase, required for uracil biosynthesis, converts 5-FOA to 5-fluorouracil, a toxic compound (yeast) |

# 1 Introduction

## 1.1 Overview of Translation Initiation

Deoxyribonucleic acid (DNA) is the molecule that contains the genetic instructions for life, but it requires complex cellular machinery to carry out those instructions. The ability to read the language of the blueprint encoded in DNA requires molecular processes and machinery by which the DNA is first transcribed into messenger ribonucleic acid (mRNA) and then translated into proteins, which carry out much of the work of the cell. The expression patterns of these proteins – which of them will be made, when, and how much of each will be made – determine the identity and function of a cell. They differentiate a neuron from a muscle cell, a dividing stem cell from a senescent or autophagic cell, and a healthy cell from a cancerous one. Gene expression is a dynamic process which also allows cells to respond to environmental conditions in real time, mounting defenses to infection and responding to stressors. In the decades since Watson and Crick's publication of the double helix, much research has been focused on the regulation of transcription of DNA to mRNA, and this is commonly referred to as 'gene expression', but it tells only part of the story. Expression of the genes encoded in DNA relies on many layers of regulation that impact not only transcription, but also mRNA stability, translation, and protein stability and localization. These are not static processes, but dynamic ones that alter protein expression, and therefore cellular identity.

Translation is the process by which mRNAs direct the synthesis of proteins in the cell. Cellular machinery called the ribosome, consisting of a small and a large subunit, carries out this process. Translation can be divided into four stages: initiation, elongation, termination, and recycling. Initiation involves delivery of the ribosome and methionyl initiator tRNA to the mRNA in preparation for elongation, as well as the crucial step of identifying an appropriate starting point on an mRNA. In the elongation step, codons are successively decoded by base-pairing to the anticodon of the appropriate tRNA, resulting in peptide bond formation and extension of a nascent protein chain. When the ribosome reaches a stop codon, the process terminates through the binding of termination factors and cleavage of the newly made protein from the last decoding tRNA. The ribosomal subunits are then released from the mRNA during recycling to be reused in future cycles of translation. Regulation occurs at each step of this process, but none has so large an effect on the proteins produced as initiation, which is generally accepted to be the rate-limiting step of translation. It is at this stage that a start codon is selected, and this process will determine the identity and quantity of proteins produced, as well as setting the reading frame. Thus, translation initiation plays an outsized role in determining the proteome of the cell.

### 1.1.1 Translation Initiation in Prokaryotes

Prokaryotic translation is relatively simple compared to the process of protein synthesis in eukaryotes. The ribosome is smaller, fewer initiation factors are required, and the start codon is identified by a defined mRNA sequence. Translation initiation in bacteria occurs cotranscriptionally, with the ribosome loading on the ribosome binding site of the mRNA as soon as the latter is made by RNA polymerase, and the ribosome physically interacts with RNA

polymerase (1). In bacteria, the initiator methionyl tRNA is formylated (fMet-$tRNA_i$) (2). Many prokaryotic mRNAs are polycistronic, containing more than one coding sequence (3). The goal of the initiation phase is to assemble the 70S bacterial ribosome (consisting of 30S and 50S subunits) at the correct start site in the translation initiation region (TIR) of an mRNA with fMet-$tRNA_i$ in the peptidyl (P) decoding site of the 30S subunit, with the assistance of the initiation factors IF1, IF2, and IF3. Initiation is the only stage of translation in which the incoming tRNA binds to the P site rather than the A site. IF1 binds to the ribosomal acceptor (A) decoding site, preventing the binding of elongator tRNAs there prior to initiation, and stabilizes the binding of IF2 and IF3 to the 30S subunit (4, 5). IF2, a multidomain GTPase, serves to recruit fMet-$tRNA_i$ to the ribosome and regulate large subunit joining (6). This factor is unusual among GTP-binding proteins because its low affinity for both GTP and GDP allows for spontaneous nucleotide exchange (7). Although IF2- fMet-$tRNA_i$ complexes do form, they are not required and do not bind to the 30S ribosome directly in this form. Thus, while IF2 assists with fMet-$tRNA_i$ recruitment, it does not serve as a vehicle for fMet-$tRNA_i$ delivery, and GTP is not required for the binding of IF2 to fMet-$tRNA_i$ (8). IF3 interferes with ribosomal subunit association, helping to ensure the accuracy of initiation, and also discriminates against mRNAs with unfavorable TIRs (9, 10). IF1, IF2, IF3, and fMet-$tRNA_i$ bind the ribosomal small subunit (SSU) to form a 30S preinitiation complex (PIC). Base pairing of the Shine-Dalgarno sequence, located 8 – 10 nucleotides upstream of the start codon, to the anti-Shine-Dalgarno sequence at the 3’ end of the 16S rRNA guides the 30S ribosome to bind mRNA with the correct start codon near the P site. The initiation factors and the binding of fMet-$tRNA_i$ facilitate precise positioning of the start codon in the P site. The 50S large ribosomal subunit then joins the 30S IC to form a 70S IC, ready to translate the selected mRNA (reviewed in (11)).

### 1.1.2 Translation Initiation in Eukaryotes

By contrast, in eukaryotes, translation proceeds through a scanning mechanism (Fig. 1.7.1). In this process, the interaction of the Shine-Dalgarno sequence with the 18S rRNA is absent, and instead each nucleotide is sequentially inspected for complementarity to the anticodon of methionyl initiator tRNA (Met-tRNA$_i$). In addition, the small 40S subunit of the ribosome associates with eukaryotic initiation factors (eIFs) including eIF1, eIF1A, eIF3, and eIF5 as well as a pre-formed ternary complex (TC) composed of GTP-bound eIF2 and Met-tRNA$_i$ to form a 43S pre-initiation complex (PIC) prior to association with mRNA. A capped mRNA is bound by the eIF4F complex, consisting of eIF4A, eIF4E, and eIF4G. eIF4E binds to the 7-methylguanylate ($m^7G$) cap of mRNA and to the scaffolding protein, eIF4G. eIF4G associates with poly-(A) binding protein (PABP), which brings the 3' poly-(A) tail of mRNA near the capped 5' end, forming a closed-loop structure. The 43S PIC is then recruited to the 5' end of the mRNA, where it binds, forming a 48S structure, and scans in the 5'-to-3' direction until an appropriate start codon is encountered. While hydrolysis of GTP on eIF2 may occur during scanning, inorganic phosphate ($P_i$) is not released. Recognition of an AUG in good context by the PIC triggers a cascade of events including release of eIF1, release of the $P_i$ from the GDP•$P_i$ on eIF2, and a conformational change of the ribosomal small subunit from an open, scanning-conducive state to a closed state in which the tRNA is bound tightly in the P site. The result of this process is Met-tRNA$_i$ base-paired to an AUG in the P site and an empty A site, in preparation for elongation (reviewed in (12)).

### 1.1.3 Alternative Models of Translation Initiation

The pathway of scanning initiation described above is the most common and best understood mechanism of translation initiation in eukaryotes, but translation can also occur via a cap-independent pathway (reviewed in (13)). Internal ribosome entry sites (IRESs), which guide binding of the ribosome directly to an open reading frame (ORF) independent of the 5' mRNA cap, have long been known to facilitate translation of viral genomes (14). Current evidence indicates that translation of several cellular eukaryotic mRNAs is initiated by an IRES, most notably GRP78/BiP, while high-throughput and bioinformatic techniques have identified thousands of putative cellular IRESs, including many in eukaryotic circular mRNAs (15-22). Other mRNA elements involved in non-canonical translation initiation include cap-independent translation elements (CITEs), translation initiator of short 5'-UTR (TISU) elements, and post-transcriptional modification with $N^6$-methyladenosine ($m^6A$) (23). Repeat-associated non-AUG (RAN) translation is a non-canonical process of translation initiation in nucleotide repeat expansions associated with neurological diseases (reviewed in (24)). There is evidence that translation may proceed via other alternative mechanisms, such as utilizing eIF2A or eIF2D for initiation at non-AUG codons under certain circumstances, although these instances are not yet well understood (25). Alternative translation initiation pathways have been shown to be used under conditions of cellular stress, including hypoxia, apoptosis, starvation, and viral infection, when cap-dependent translation is globally down-regulated (13, 23). Thus, these mechanisms may play a vital role in the cellular response to stress and disease.

## 1.2 Translation Initiation in Human Health

Given its role in determining the identity and abundance of cellular proteins, translation initiation unsurprisingly plays an important role in the maintenance of human health. Start codon selection plays a role in determining and maintaining cell fate throughout life, but it is particularly important in instances that require rapid changes in gene expression. During early embryonic development, in the absence of transcription, translational control on maternal mRNAs is responsible for the patterning of cell fates and establishing the embryonic axis (26). Much of this control is accomplished through silencing of mRNAs by eIF4E binding proteins (4E-BPs) that prevent initiation by competing with eIF4G for binding to the cap binding protein eIF4E, reducing eIF4F assembly. (27, 28). Translation initiation also appears to be of particular importance in neurobiology. Local translation of mRNAs at synapses plays an important role in learning, memory, and synaptic plasticity (29, 30). There is evidence that eIF2α phosphorylation, and attendant translational upregulation of *ATF4* mRNA, encoding a stress-responsive transcription factor, may suppress memory formation (28, 31).

Accordingly, dysregulation of translation initiation is involved in many neurological disorders (32). Fragile X syndrome, a monogenic disorder of neurodevelopment is caused by an increase in CGG repeats in the FMRP gene, resulting in inhibition of translation initiation through binding to CYFIP1, a 4E-BP (33, 34). Leukoencephalopathy with vanishing white matter, an autosomal recessive leukodystrophy, results from mutations in any of the five subunits of eIF2B, the guanine nucleotide exchange factor for eIF2 (35, 36). MEHMO syndrome, an X-linked intellectual disability, is caused by mutations in the *EIF2S3* gene, encoding eIF2γ, that impair

ternary complex formation (37). In addition to these genetic disorders, initiation of protein synthesis is also implicated in many neurodegenerative diseases. There is evidence that translation is altered in the cerebral cortex of patients with Alzheimer's disease (AD), including altered expression of eIF2α, eIF3 subunits, and eIF5 (38). Mutations of several proteins involved in translation initiation are known to cause Parkinson's disease (PD), most notably *EIF4G1*, encoding eIF4G (39, 40). The activation of mammalian target of rapamycin (mTOR), a central regulator of protein synthesis which stimulates cap-dependent translation initiation through phosphorylation of 4E-BP1 and p70 S6 kinase (S6K), has been implicated in a number of neurological diseases. The brains of patients with AD and PD show evidence of altered mTOR and eIF2 signaling, including phosphorylation of eIF2α, mTOR, S6K, eIF4E, and 4E-BP1 (41-43).

Protein synthesis is one of the most energetically demanding processes in the cell, thus requiring cells to regulate translation in response to their metabolic state. Phosphorylation of eIF2α, discussed in Section 1.5.2, is a major regulator of translation in response to nutrient stress, and dysregulation of this process, such as through mutations in eIF2α kinases, 4E-BPs, or S6Ks, can result in metabolic defects (28). In particular, mutations that affect PERK, an eIF2α kinase involved in the cellular response to ER stress, or its ability to phosphorylate eIF2α result in diabetes and glycemic disorders (44). Wolcott-Rallison syndrome, an autosomal recessive disorder that results in infant or early childhood onset of diabetes as well as neurological defects and hypothyroidism, is caused by a mutation in PERK (45). Insulin stimulates phosphorylation of 4E-BP1, disabling its ability to inhibit eIF4F assembly, and activates eIF2B, which also responds to diet and exercise (46-48). S6K1, a serine-threonine protein kinase, acts downstream

of mTOR to phosphorylate multiple targets including Rps6 and eIF4B, leading to activation of translation and ribosome biogenesis when nutrients are plentiful. It has been shown to be required for development of mouse pancreatic β-cells, and mice lacking S6K1 display reduced insulin production, leading to hyperglycemia upon glucose challenge (49). Thus, regulators of translation such as PERK, 4E-BP1, and S6K1 present attractive targets for new diabetes therapies (50).

Upregulation of the expression or activity of initiation factors correlates with cancer and heart disease, presumably by dysregulating cell growth and proliferation (28). In fact, alterations in the process of translation underlie changes in many cellular processes involved in cancer progression, including metabolism, cell growth, cell cycle progression, tumorigenesis, and metastasis (51, 52). The ribosome acts as a hub for regulation in cancer cells, receiving signals from oncogenic pathways involving the RAS or MYC oncoproteins and PI3K, which control the synthesis and function of many factors involved in translation initiation in order to promote tumor growth (53). For example, gene amplification and overexpression of eIF4E were observed in breast cancer (54). Overexpression of eIF4E results in tumorigenesis in cell culture (NIH 3T3, Rat2 fibroblasts, HeLa cells), while administration of antisense oligonucleotide to eIF4E severely retards tumor growth in mice (55-57). The cap binding protein also plays an important role in cardiac health. Increased cardiac load results in hypertrophic growth of the heart muscle, which is coupled to an increase in protein synthesis through upregulation of eIF4F components and phosphorylation of eIF4E (58, 59). Following treatment with a small molecule inhibitor of sphingosine-1-phosphate (S1P) lyase, a stress activated enzyme that contributes to oxidative stress through S1P catabolism, an increase in eIF4E phosphorylation correlated with other anti-

ischemic biomarkers (60). Additionally, changes in mTOR activity during hypertrophy and heart failure modulate translation through altered 4E-BP1 binding (61).

Mutations in the translational machinery are also known to cause several inherited diseases (reviewed in (62)). These include the neurological disorders mentioned above, such as Fragile X and MEHMO syndromes. They also include a class of disorders arising from mutations in ribosomal proteins or key ribosome biogenesis factors, known as ribosomopathies (63). Although these diseases presumably affect ribosomes throughout the body, they display surprisingly tissue-specific phenotypes, many involving bone marrow impairment (e.g., Diamond-Blackfan anemia, 5q- syndrome)(64). The reasons for this specificity of symptoms remain elusive. It has been suggested that they may result from a preference of "specialized ribosomes" lacking a particular ribosomal protein for a subset of mRNAs, while an alternative model asserts that changes in ribosome abundance broadly reprograms translational efficiencies and certain tissues are more sensitive than others to this re-programming (63, 65).

Other disorders result from mutations in the 5'-untranslated region (UTR) of a gene. Hyperferritinemia, an increase in the iron-storage protein ferritin that leads to cataracts, results from a mutation in a "stem-loop" secondary structure in the 5'UTR of ferritin mRNA that mediates translational repression by an RNA binding protein, IRE-binding protein (IRP), in iron-depleted cells (66). As a result of ribosomal scanning in eukaryotes, the presence of an upstream open reading frame (uORF) frequently reduces translation at the primary start codon downstream. This is because the scanning PIC encounters the AUG of the uORF first, and reinitiation following termination of uORF translation is typically inefficient. Thrombocythemia,

or excessive platelet production, results from the loss of a uORF in the thrombopoietin gene (67). On the other hand, a mutation in the gene CDKN2A, encoding p16, introduces an inhibitory uORF, resulting in lowered levels of this tumor suppressor and increased incidence of melanoma (68). Thus, it is apparent that a thorough understanding of translation initiation would have a dramatic impact on the understanding and treatment of a broad swath of human disease.

### 1.3 The Ribosome

An understanding of the structure and function of the ribosome is critical in understanding protein synthesis. The ribosome and associated translation factors form the translational machinery of the cell. Although they were observed by dark-field microscopy in the 1930s (69) and by electron microscopy in the 1950s (70), the ribosome was not confirmed as the site of protein synthesis until Paul Zamecnik's 1955 experiment demonstrating that [$^{14}$C]-labeled amino acids are transiently associated with ribosomes before their incorporation into proteins (71, 72). Further research has cemented the ribosome as a nucleoprotein complex that does the work of reading the codons of mRNA and synthesizing the encoded protein. As a result of this vital and ancient function, the ribosome is relatively well conserved throughout evolution. Here, we will discuss the structure of the ribosome, which informs its function, the importance of the ribosomal proteins, and how ribosomes are assembled within the cell.

### 1.3.1 Ribosomal Structure

Since the first low (~50 Å) resolution micrographs of "microsomes", advances in X-ray crystallography and, more recently, cryo-electron microscopy (cryo-EM) have allowed more and more detailed studies of ribosomal structure on the scale of 2.5 – 3.5 Å. In all organisms, the ribosome consists of a small and a large subunit (LSU), each composed of rRNA and ribosomal proteins. The small subunit (SSU) can be divided into head and body regions, connected by a narrow neck, with a relatively flat solvent face and an intersubunit face, which contains the A, P, and E decoding sites (Fig. 1.7.2 A-B). Viewing the SSU from the interface side, the body of the SSU contains a large platform on the right and a smaller shoulder to the left (Fig. 1.7.2 C). The mRNA enters the ribosome through an entry channel above the shoulder and passes through a groove between the head and body of the SSU before exiting through an exit channel above the platform (Fig. 1.7.2 C). On this path, mRNA passes through the A (aminoacyl), P (peptidyl), and E (exit) tRNA binding sites, which are formed at the interface of the SSU and LSU. The SSU contains the portions of the decoding sites where the anticodon stem-loops of the tRNAs bind and base-pair with their cognate codons in mRNA. The LSU is rounded, and features a central protuberance and two stalks (L1 and L7/L12) (Fig. 1.7.2 D). The LSU also contains the peptidyltransferase center (PTC), which catalyzes peptide bond formation, and the exit tunnel through which the nascent peptide chain exits the ribosome.

Ribosomal subunits are described in Svedberg units (S), a measure of a particle's sedimentation rate during ultracentrifugation. Because these units depend on the mass, density, and shape of a particle, they are not additive. In bacteria, the 30S small subunit, composed of 16S rRNA and 21

ribosomal proteins, and the 50S large subunit, made up of a 5S rRNA, a 23S rRNA, and 31 ribosomal proteins, form a 70S bacterial ribosome that is ~250 Å in diameter. In contrast, the eukaryotic ribosome is larger (80S), with a 40S SSU comprising 18S rRNA and 33 ribosomal proteins and a 60S LSU consisting of a 28S rRNA, a 5.8S rRNA, a 5S rRNA, and 49 ribosomal proteins. In spite of these differences, the core structure of the ribosome is quite well conserved throughout all kingdoms. Rather than disturbing this core structure, most of the additional rRNA in the eukaryotic ribosome is present in discrete expansion segments (ES) or variable regions (VR). There are five ESs and five VRs in the SSU as well as 16 ESs and 2 VRs in the LSU. Most of these additions are present in the foot of the SSU and on the back and sides of the LSU, leaving the subunit interfaces relatively unchanged. The increased complexity of the eukaryotic ribosome affords higher organisms more complex regulatory control of translation, including a more nuanced mechanism of translation initiation (reviewed in (73, 74)).

The first crystal structures of the ribosome were resolved in 2000 (75-77) and subsequent crystal structures revealed the architecture of bacterial ribosomal particles (78-81). The first cryo-EM structure of the eukaryotic ribosome was modeled with heavy reliance on existing bacterial and archaeal structures (82). Subsequent eukaryotic cryo-EM structures shed light on specific interactions and extended mapping of the ribosome at a resolution of 15 – 5.5 Å (83-89). Higher resolution (3.0 – 3.9 Å) crystal structures allowed mapping of all r-proteins in the 80S ribosomes of Tetrahymena (90, 91) and *S. cerevisiae* (92). A growing number of structures of the eukaryotic ribosome have been published in the last decade, and advances in cryo-EM that allow visualization of multiple states at high resolution within a single sample have resulted in tremendous advances in our understanding of eukaryotic ribosomes in the last few years.

Structures of the 80S ribosome in dogs, pigs, and humans have been solved (93-95). Several other structures illuminated the mechanisms of viral IRESs (96-103). A number of ribosomal structures from pathogenic protozoans have been solved, providing information that may lead to interventions for the diseases caused by these common microorganisms (104-108). Other recent cryo-EM structures have shed new light on the process of translation initiation, as discussed further below (109-116). While structural data have allowed many insights into the function of the ribosome, it is important to remember that no structure can tell the whole story. It is difficult to image the many transient states through which the ribosome passes during each stage of translation, and even more difficult to assign these images to their respective stages in the process. Much genetic, biochemical, and structural work remains to fully understand this molecular machine and its roles in protein synthesis.

To further our discussion of eukaryotic translation initiation, we will focus on the structure of the *S. cerevisiae* 40S small ribosomal subunit. As discussed above, the SSU can be divided into the head, with a beak on the entry channel side, and the body, with a large platform on the exit channel side (right), a smaller shoulder on the entry channel side (left), and a right and left foot. In yeast, the entry and exit channels, formed by a groove between the SSU head and body that extends from the intersubunit face to the solvent face with pores opening on the solvent side, each hold 12 nucleotides of mRNA. In the entry channel, a latch between helices 18 and 34 opens to allow mRNA loading and scanning. Between the entry and exit channels, the mRNA traverses the neck of the 40S subunit, passing through the 30S portions of the A, P, and E decoding sites. The rRNA in all three sites is largely conserved among bacteria, archaea, and eukaryotes, suggesting that the decoding mechanism is also conserved. However, as elsewhere in

the 40S subunit, increased structural complexity allows for greater regulation in eukaryotes, and additional contacts between ribosomal proteins and the tRNAs may allow another level of discrimination against incorrect tRNAs or an incorrect codon:anticodon duplex.

During translation initiation, the SSU binds to and scans along the mRNA as a 43S pre-initiation complex (PIC) made up of the 40S subunit, a ternary complex of methionyl-initiator tRNA and GTP-bound eIF2, and initiation factors eIF1, eIF1A, eIF5, and the multi-subunit eIF3. The TC binds in a position that is roughly perpendicular to the 40S intersubunit face, with the anticodon stem loop of Met-tRNA$_i$ bound in the P site. eIF1 and eIF1A bind the 40S on the intersubunit face near the P site, with the globular domain of eIF1A occupying the A site. Much of eIF3 binds on the solvent face of the SSU, with some subunits reaching the mRNA entry and exit channels. The positions of the eIF5 C-terminal domain and several subunits of eIF3 remain the subject of continued study. Upon recognizing an AUG start codon, the ribosome undergoes a series of structural changes that include dissociation of eIF1 and release of $P_i$ from the GDP•$P_i$ bound to eIF2. These changes include a shift from a widened mRNA channel, an open latch, and a widened P site to a closed latch and a constricted mRNA entry channel and P site, facilitating more extensive contacts between the 40S subunit and Met-tRNA$_i$, as well as a tighter binding of Met-tRNA$_i$ to the PIC in the closed state ($P_{IN}$) (Fig. 1.7.3 A-B). These alterations of the PIC are accompanied by a downward movement of the 40S head and a 7 – 8 Å change in the pitch of h28, compressing the neck of the 40S (Fig. 1.7.3 A) (reviewed in (117)).

While it is useful to think of the ribosome as being in either an open, scanning-conducive state or a closed, scanning-incompetent state, it is likely that many individual interactions continually

modulate the ribosomal structure in addition to the large movement of the head relative to the body and latch closure. It is important to remember that the ribosome is not a static scaffold on which translation takes place. Rather, it is highly dynamic, alternating swiftly through many different structural states to bring together the correct components and provide the correct environment to facilitate synthesis of the correct proteins in the correct quantity at the correct time. The inclusion of ribosomal proteins in various signaling pathways, such as S6, which stretches from the right to the left foot of the 40S subunit and is phosphorylated in response to mTOR signaling, suggests that the ribosome is subject to dynamic regulation in response to cellular conditions in addition to the dynamism required to accommodate protein synthesis.

### 1.3.2 Ribosomal Proteins in the Eukaryotic Ribosome

Ribosomal proteins make up ~40% of the ribosome by mass, and are among the best conserved of all proteins. Of the 33 ribosomal proteins of the SSU, 15 have bacterial homologs, and an additional 12 have homologs among archaea, while 6 r-proteins are specific to eukaryotes. Beyond containing 18 r-proteins that are absent in bacteria, 9 of the conserved proteins contain extensions specific to eukaryotes, resulting in a much lower ratio of rRNA:protein approaching 1:1 (91). The bulk of this additional protein mass is present on the solvent faces of the ribosomal subunits, although some tail extensions form new long-range contacts within the LSU or SSU. Historically, ribosomal proteins from different species were named independently, resulting in the same number being assigned to unrelated proteins. Here, we use the system described in Ban, et al., 2014, in which the protein number is accompanied by a prefix designating its kingdom to avoid confusion. Proteins specific to bacteria are designated with a b (bacterial) and eukaryotic

ribosomal proteins without bacterial homologs are specified with an e (eukaryotic), while proteins found in bacteria, archaea, and eukaryotes are designated with a u (for universal). Similarly, r-proteins specific to archaea would be designated with an a (archaeal), although none have been found (118).

In the SSU, three eukaryote-specific r-proteins are present in the beak region (eS10, eS12, and eS31), and these appear to take the place of a missing section of helix 33 (h33), which is shortened in eukaryotes. Many r-proteins unique to eukaryotes interact with expansion segments in the rRNA, which are likewise specific to eukaryotes (eS4, eS6, eS7, eS8) and presumably act to stabilize this additional rRNA architecture. At the entry channel, eS30 and a eukaryote-specific extension of uS3 may assist with the resolution of mRNA secondary structure. This is a vital function, since mRNA must be single stranded to pass through the entry and exit channel pores. In eukaryotes, uS3 has also been implicated in stabilization of the closed PIC through binding to mRNA (119), and contains a long C-terminal extension that interacts with scaffold protein RACK1 on the 40S solvent face, representing a possible mechanism for signal transduction pathways that target RACK1 to influence translation, as well as a link to ribosomal quality control. An important difference in eukaryotic ribosomes is the lack of the anti-Shine-Dalgarno sequence in rRNA and its Watson-Crick base-pairing with mRNA upstream of the start codon. Instead, eS26 and eS28 surround the 3' end of the 18S rRNA, overlapping the binding site of bS21. The β-barrel of eS28 is analogous to that of bS1. eS26, discussed further in Chapter 4 of this work, locks the 3' end of the 18S rRNA in place, while it is flexible in bacteria. These differences clearly reflect the adaptation of the eukaryotic ribosome for translation initiation via

mRNA scanning. In general, the additional r-proteins and extensions in eukaryotes provide additional opportunities for regulation of translation in these organisms.

### 1.4 Evidence Supporting the Current Model of Translation Initiation

As discussed above, translation initiation in eukaryotes generally occurs via a scanning mechanism. Unlike the direct base-pairing to the Shine-Dalgarno sequence utilized in bacteria to place AUG in the P site, start codon recognition in eukaryotes is a much more complicated process, and one that allows for much more nuanced regulation. Specifically, in eukaryotes, (i) Met-tRNA$_i$ does not bind to the SSU independently, but must be assembled into the TC and requires the assistance of several initiation factors for rapid incorporation into the 43S PIC, (ii) mRNA recruitment by the PIC requires recognition of a 5' 7-methylguanosine cap on the mRNA, (iii) eukaryotic mRNAs generally have longer 5' UTRs, many of which contain stable secondary structures which must be unfolded to allow ribosomes to traverse them, and (iv) scanning ribosomes must recognize an appropriate start codon based on its identity and sequence context without the benefit of being positioned in the P site via base-pairing with the rRNA (120).

The GTPase eIF2, in its GTP-bound form, recruits Met-tRNA$_i$ to the ribosome by forming the TC. A multifactor complex (MFC) of eIF1, eIF1A, eIF3, and eIF5 promotes TC recruitment. In addition to binding to one another, these factors also bind directly to the 40S ribosome along with TC, forming a highly stable 43S PIC. eIF1 and eIF1A promote the open conformation of the ribosome, facilitating both rapid TC loading and mRNA recruitment. Binding of the PIC to

mRNA near the 5’ cap is also stimulated by interactions between eIF3 and eIF5 in the PIC and the eIF4F complex (composed of eIF4G, eIF4E, and eIF4A) bound to the cap, and by poly(A)-binding protein (PABP) and RNA helicases. The PIC then scans the mRNA base-by-base in a 5’-to-3’ direction until an AUG in appropriate sequence context is recognized. The GTP on eIF2 may be hydrolyzed during scanning, but release of $P_i$ from eIF2-GDP-$P_i$ is blocked by the scanning PIC until an appropriate start codon is recognized. Base pairing between the AUG and the anticodon of Met-$tRNA_i$ in the P site triggers the cessation of scanning, dissociation of eIF1, release of $P_i$ from eIF2-GDP-$P_i$, dissociation of eIF2-GDP from Met-$tRNA_i$ in a complex with eIF5, followed by release of other initiation factors, and culminating in joining of the large (60S) subunit to the 40S initiation complex, catalyzed by eIF5B. The end result of translation initiation is an 80S ribosome containing Met-$tRNA_i$ base paired to an AUG in the P site, ready to proceed into the elongation phase of translation.

### 1.4.1 TC formation

The initiator tRNA required for recognition of an AUG start codon is delivered to the 40S ribosome in ternary complex with GTP-bound eIF2. The heterotrimeric eIF2 complex acts as a G-protein, requiring eIF5 as a GTPase activating protein (GAP) and eIF2B as a guanine nucleotide exchange factor (GEF), and displays 100-fold greater affinity for $tRNA_i$ when bound to GTP vs. GDP (121-123). The loss of the positive interaction between eIF2 and Met-$tRNA_i$ upon GTP hydrolysis facilitates release of eIF2 for reuse following start codon recognition (123). A prominent mechanism of translational control in eukaryotes operates through phosphorylation of eIF2, limiting the availability of eIF2•GTP to form new TC by inhibition of eIF2B. Many

aspects of eIF2 are discussed in greater detail in section 1.5. Conserved features of Met-tRNA$_i$ contribute to TC formation, including the bound methionine. eIF2 displays > 10-fold lower affinity for deacylated tRNA$_i$ and 20-fold lower affinity for elongator methionyl tRNA (Met-tRNA$_e$) compared to Met-tRNA$_i$ (123). These affinity differences likely serve as quality checks to ensure that only tRNA$_i$ charged with methionine is loaded onto the PIC. While initiator Met-tRNA$_i$ can replace the elongator Met-tRNA$_e$ in the elongation phase of protein synthesis *in vivo*, Met-tRNA$_e$ cannot act in initiation (124). The first base pair of the tRNA$_i$ acceptor stem, A1:U72, imparts preferential eIF2 binding to Met-tRNA$_i$ versus Met-tRNA$_e$, where it is replaced by a G:C base pair (124, 125). This base pair has also been shown to improve Met-tRNA$_i$ binding to eIF2 by orienting the methionine moiety in its recognition pocket (123). The G31:C39 base pair in the ASL cooperates with A54 and A60 in the T loop to impart initiation-specificity to Met-tRNA$_i$ by strengthening the hydrogen bond network, allowing appropriate deformation of the anticodon stem during AUG recognition, as well as preventing the use of initiator tRNA during elongation (126).

### 1.4.2 PIC assembly

Factors eIF1 and eIF1A bind directly to the 40S subunit and facilitate loading of TC (127-129). Although these factors bind to separate sites on the 40S, their 40S-binding is thermodynamically coupled, and both factors are required for stable and efficient binding of TC *in vitro* (130, 131). eIF3 also enhances recruitment of TC to the 40S subunit and stabilizes the PIC (132-136). Cooperative binding of these factors induces conformational changes in the SSU that result in

rotation of the 40S head, opening the latch and widening the mRNA entry channel (127, 137, 138).

The ability of eIF1 and eIF1A to stimulate TC binding to the 40S support a sequential model of 43S PIC assembly, but an *en masse* mechanism has also been proposed. In this latter model, factors eIF1, eIF2, eIF3, and eIF5 cooperate to form a multi-factor complex (MFC) that binds to the PIC simultaneously with the Met-tRNA$_i$ (139). This complex has been isolated in yeast, as well as plant and mammalian cells, and can be separated from the 40S subunit in cellular extracts (139-141). The MFC may enable cooperative recruitment of initiation factors to the PIC (142). It is consistent with the multiple known interactions among these factors, which are required for efficient protein synthesis *in vivo*. eIF5 has independent contacts with eIF1, eIF1A, eIF2β, and the NTD of eIF3c (143, 144). eIF1 binds directly to eIF3, as does eIF2, through the eIF3a-CTD (145-147). The exact function of the MFC and relative biological importance of an MFC pathway to PIC formation vs. the individual recruitment of initiation factors are not yet known (142).

### 1.4.3 Closed loop formation and eIF4A helicase activity stimulate PIC attachment

Scanning ribosomes typically initiate at the 5'-proximal AUG because 43S PICs load onto mRNA and begin scanning near the 5' cap. Prior to 43S PIC binding, the cap binding protein, eIF4E binds the 7-methylguanosine cap and recruits the scaffolding protein eIF4G and the RNA helicase eIF4A to the 5' end of mRNA. In addition to targeting the eIF4F complex (made up of eIF4E, eIF4G, and eIF4A) and the 43S PIC to the 5' end of mRNA, cap binding also ensures that

a bound RNA is, in fact, mRNA. Interaction with eIF4G enhances the ability of eIF4E to bind the cap (148). Activation of eIF4A by eIF4G and eIF4E results in single-stranded RNA near the 5' cap that acts as a binding site for the PIC. Although eIF4A can bind and unwind mRNA on its own, these activities are enhanced by its assembly in the eIF4F complex (142). Although there is some evidence that more structured mRNAs display a stronger requirement for eIF4A, it has been found to promote translation of all mRNAs, even those with short, unstructured leaders (149-152). More recent evidence from *in vitro* studies and ribosome profiling suggests that the majority of cellular mRNAs have a strong dependence on eIF4A, regardless of their structural complexity (153, 154). There is also evidence to suggest that eIF4A enhances mRNA recruitment to the PIC by altering the structure of the 40S subunit independent of its helicase activity (155). 48S PIC assembly is greatly enhanced by the presence of eIF4B, particularly for long mRNAs with structured 5'-UTRs, or with high GC content (152, 156, 157). eIF4B binds directly to the 40S subunit near the mRNA entry channel and modulates the ribosomal conformation in this region (158). Recent data demonstrate that eIF4B can stimulate translation independently of eIF4A (159).

In addition to binding eIF4E and directly activating eIF4A, eIF4G recruits the 43S PIC to mRNA by binding the PIC directly, through interactions with eIF5 in yeast and eIF3e in mammals (139, 152, 160-164). While eIF4E binds the 5' end of mRNA, PABP binds the poly(A) tail at the mRNA 3' end. Interactions between PABP and both eIF4E and eIF4G result in circularization of the mRNA, forming a closed-loop structure with eIF4F tightly bound to the cap, which protects the mRNA from exonucleases and may aid in the recycling of ribosomes released at the stop codon following termination to the start codon of the same mRNA (165-169). It was found that

initiation complexes can form on an AUG codon even when it occurs as the first triplet at the mRNA 5' end in a reconstituted system, indicating that mRNA is threaded into the entry channel from the 5' end rather than slotting into the mRNA binding cleft with the 5'cap-eIF4F assembly located just outside the exit channel pore, which would prevent utilization of a 5'-terminal AUG. Threading of the 5' end into the entry channel envisioned in the first mechanism would likely necessitate dissociation of eIF4E from the cap (170). However, this evidence does not rule out the possibility that PICs may load near, but downstream of, the 5' cap *in vivo*.

eIF3 plays a major role in mRNA recruitment. It promotes the binding of the 43S PIC to native mRNAs and stimulates 48S formation (130, 152, 171-174). This is consistent with its role in stabilization of bound mRNA at both the entry and exit channels (136). eIF3j was shown to interfere with mRNA binding to 40S complexes lacking TC, indicating that it may impose a requirement for a completely assembled PIC prior to mRNA recruitment (152, 175). Binding domains for the $m^7G$ cap have been identified in mammalian eIF3d and eIF3l, suggesting that these regions of eIF3 may take the place of eIF4E in binding the 5' cap during PIC attachment, consistent with the suggested dissociation of eIF4E during this process (170, 176). However, as these subunits are not present in yeast, the proposed dissociation of eIF4E from the cap to permit threading of the 5' end of the mRNA into the entry channel may be only transitory with eIF4E re-binding the cap when it emerges from the exit channel. In summary, it appears that a network of interactions among the components of the PIC and members of the eIF4F complex facilitate mRNA recruitment and act to stabilize 43S PICs on mRNA (142).

### 1.4.4 Scanning

Following mRNA recruitment, the PIC scans in a 5' – 3' direction along the mRNA. The basis for this directionality is not entirely understood, although the preference of eIF4F for an RNA duplex substrate with a 5' overhang for eIF4A helicase activity may play a role (177). It has long been known that ATP is required for scanning, presumably for the ATP-helicase activity of eIF4A (178). It is not clear that eIF4A is a processive helicase. Rather, it is possible that directionality is imparted to the scanning process through a Brownian ratchet mechanism wherein helicases remove structure 3' of the PIC and structures reform 5' of the complex (179). This hypothesis fits with evidence for limited backwards scanning (180). However, recent evidence suggests that eIF4A may act via a processive mechanism in complex with eIF4B and eIF4G (181). The rate of scanning *in vivo* has been measured at ~8 bases/s (182). Longer 5'-UTRs do not result in decreased translational efficiency, suggesting that the 5'-UTR is scanned by more than one PIC simultaneously (183). While the anticodon of Met-tRNA$_i$ is used to identify the AUG start codon, the precise method of sampling the mRNA sequence is not understood. Scanning requires the unwinding of secondary structure in order to allow the mRNA to thread through the 40S, as well as a conformation of the 40S subunit itself that is conducive to processive movement along the mRNA (12). Here, we will discuss the activity of helicases required for the dissolution of mRNA secondary structure. 40S conformation will be addressed in section 1.4.5.

While it is logical that the DEAD-box helicases that function in PIC attachment to the mRNA are also utilized to unwind secondary structure during subsequent scanning, in practice it is difficult

to differentiate these two functions. It is known that depletion of eIF4A and eIF4B inhibit translation, although they are not sufficient to unwind strong stem loops (< -19 kcal/mol) (184). Even small decreases in cellular concentrations of eIF4A reduce bulk translation *in vivo* (185). Recent work has demonstrated that eIF4A can unwind a large stem loop on RNA tethered in an optical trap, but eIF4A is nonetheless a relatively weak helicase (177, 181, 186). By contrast, DEAD-box helicases Ded1 and Dhx29 do promote unwinding of strong stem loops (184, 187). Ded1 (Ddx3 in mammals) is essential for yeast growth and stimulates bulk translation *in vivo*, as well as the translation of reporter mRNAs with long or structured 5'-UTRs (153, 183, 188, 189). Consistent with this, ribosome profiling of conditional *ded1* mutants in yeast revealed that while Ded1 is broadly required for translation, mRNAs with longer and more structured 5'-UTRs display a Ded1 hyperdependence (153). Further, Ded1 was shown to promote *in vitro* recruitment of PICs and assembly of 48S complexes on these hyperdependent mRNAs and synthetic mRNAs containing stem loops in their 5'-UTRs in complex with eIF4F (190).

There is evidence that the DExH-box protein Dhx29 is also required to translate highly structured mRNAs. Dhx29 binds to the 40S subunit and hydrolyzes nucleotide triphosphates, stimulated by the 43S PIC (187). It promotes translation initiation *in vivo* and is required for 48S PIC assembly on mammalian mRNAs with structured 5'-UTRs *in vitro* (187, 191). Recent cryo-EM structures show Dhx29 bound on the shoulder of the 40S subunit, near the mRNA entry channel latch (109). This site coincides with the binding site of eIF3j, indicating that these two factors do not bind in tandem (175). Importantly, Dhx29 was found to stabilize the PIC, facilitating visualization of the mammalian ribosome by cryo-EM (73, 109). While Dhx29 is not present in yeast, it enhances translation of many mRNAs in mammals.

### 1.4.5 Start codon selection

While the scanning PIC generally initiates at the first 5' proximal AUG, this is not always the case. For example, it is known to skip over AUG codons that are less than 5 nucleotides from the 5' cap, or those in poor sequence context. Accurate translation initiation thus depends on the ability of the ribosome to scan past some AUG codons and reliably initiate at others. In addition, certain start codons, such as near cognate codons (e.g., UUG) or AUG codons in poor context, act to regulate the levels of the proteins that they encode. Such a strategy is employed in the autoregulation of eIF1 described below. In the absence of a defined PIC-positioning sequence such as the bacterial Shine-Dalgarno sequence, how does the eukaryotic ribosome identify the correct start codon for initiation? Start codon recognition in eukaryotes is a complex process that involves base-pairing of the anticodon of Met-tRNA$_i$ to the start codon, conformational changes in the 40S subunit, and a web of molecular interactions among mRNA, tRNA, and rRNA sequences, ribosomal proteins, and initiation factors.

Base pairing between Met-tRNA$_i$ and AUG was established as a primary determinant of start codon selection through yeast genetic experiments. Cells in which the *HIS4* start codon was mutated from AUG to AGG were unable to synthesize His4 and were His$^-$, but became His$^+$ on ectopic expression of a Met-tRNA$_i$ variant in which the anticodon was mutated to CCU from native CAU. In these cells, introduction of an AGG codon upstream and out of frame from the *HIS4* coding sequence (CDS) inhibited His4 production (192). Base pairing between AUG and Met-tRNA$_i$ in reconstituted mammalian PICs stabilizes the binding of TC to the 40S subunit (12). In a reconstituted yeast system, the on-rate of TC binding to 48S PICs was dramatically lowered

when the second or third position of the AUG start codon was altered, a defect that was suppressed by compensatory mutations in the Met-tRNA$_i$ anticodon (193). That the TC off-rate was less affected by the absence of an AUG start codon in these experiments (a four-fold range, compared to 70-fold change in on-rate) suggests a conformational change in which Met-tRNA$_i$ is more tightly bound following start codon recognition (193, 194).

The conformational change suggested by Kolitz, et al. likely corresponds to the shift of the PIC from an open to a closed state upon AUG recognition (Fig. 1.7.4)(12). The binding of factors eIF1 and eIF1A was found to promote an open conformation of the yeast 40S subunit in which contacts between rRNA residues of h34 in the 40S head and h18 in the body, dubbed the 'latch', were lacking (117). This conformation was presumed to promote scanning of the mRNA. Subsequent crystal structures of 40S•eIF1•eIF1A complexes from Tetrahymena and mammals displayed a closed latch, however (138, 195). A recent cryo-EM structure of a partial yeast 48S complex containing eIF1 and eIF1A obtained using mRNA with an AUC start codon and WT eIF2 and Met-tRNA$_i$ (rather than mutants expected to stabilize the closed complex) displayed an open latch (111). This structure (py48S-open) revealed an upward movement of the 40S head in response to a 7 – 8 Å change in the pitch of h28, resulting in an open latch and widened mRNA entry channel, as well as a widened P site. A second complex, also containing eIF1 and eIF1A, but utilizing an mRNA with an AUG start codon, demonstrated a closed latch similar to previous py48S structures (py48S-closed), suggesting that eIF1 and eIF1A are not sufficient for stabilization of the open complex (117). In the py48S-open complex, there are significantly fewer contacts between the PIC and mRNA in the entry channel, consistent with the idea that this conformation is conducive to mRNA attachment and scanning. This structure also lacks

many canonical contacts between Met-tRNA$_i$ and the 40S body in the P site as a result of an ~7 Å lateral displacement of the 40S head and Met-tRNA$_i$ ASL. It has been proposed that Met-tRNA$_i$ is not fully engaged in the P site in the open conformation of the PIC ($P_{OUT}$ state) (196). This conformation is not compatible with start codon recognition, leading to the hypothesis that Met-tRNA$_i$ moves rapidly between the $P_{OUT}$ and $P_{IN}$ states during scanning in order to sample mRNA codons (12). Met-tRNA$_i$ in the py48S-open complex is not in the $P_{OUT}$ state and is instead base-paired to the AUC codon. However, the reduced contacts between the ASL and the 40S body may facilitate reiterative sampling of mRNA triplets during scanning (117). By contrast, the py48S-closed structure displays a constricted mRNA entry channel and fully formed P site that presumably stabilize TC binding in the closed state and facilitate the cessation of scanning following start codon selection (111).

Significant evidence exists that rRNA residues of the 40S subunit are involved in the recognition of an appropriate start codon. Genetic analysis of yeast 18S rRNA identified residues of h28 and h44 involved in stable binding of Met-tRNA$_i$ to the PIC, particularly the A928:U1389 base pair and conserved contacts with the P site codon or tRNA$_i$ (197). The A1193U substitution in h31 in the P site impaired start codon recognition *in vivo* (198). Evidence indicates that minor groove interactions of 18S residues G1575 and A1576 in h29 with highly conserved G-C base pairs in the anticodon stem loop (ASL) may stabilize Met-tRNA$_i$ binding and prevent the binding of alternative tRNAs as well as discriminating against non-AUG codons (12). Structural data support a role for h29, along with h24, in locking tRNA$_i$ into the P site, and increased sensitivity of G1575 and A1576 to hydroxyl radical cleavage in the presence of an AUC start codon rather than an AUG is consistent with this interpretation (110, 195, 199). Indeed, many rRNA residues

in the P site and mRNA entry channel displayed reduced cleavage in PICs reconstituted with an AUG versus AUC mRNA, suggesting that these residues are less accessible following start codon recognition and transition to the closed PIC conformation and $P_{IN}$ state (Fig. 1.7.4) (199). Structural data indicate that progressive rotation of the 40S head brings residues of the 18S rRNA into direct contact with $tRNA_i$ in the P site (110, 111, 117, 195). These detailed structures have allowed better predictions of the effects of specific residues in the P site on start codon selection, but the difficulty of creating and maintaining mutations in 18S rRNA has delayed further research in this interesting area.

Specific interactions of Met-$tRNA_i$ with the PIC beyond base-pairing to the start codon also play an important role in allowing recognition of an AUG. eIF1 and the eIF1A CTT both impede rearrangement of Met-$tRNA_i$ to its closed state position through steric and electrostatic interactions with $tRNA_i$ itself, as discussed in further detail below (199, 200). The A1:U72 base pair, whose role in TC assembly was discussed above, also plays a role in stable binding of TC to the PIC (126). In addition, three consecutive G:C pairs in the ASL known to promote binding of Met-$tRNA_i$ to the P site in bacteria and to enhance the stability of reconstituted mammalian PICs have recently been demonstrated to modulate start codon selection (10, 201-203). Disruption of the third G:C pair, G31:C39, produced hyperaccuracy phenotypes, conferring increased selection of near-cognate UUG codons, and these phenotypes were abolished by substitutions of A54 and A60 in the T loop (203). Transition to the closed conformation of the PIC requires deformation of $tRNA_i$, and it is possible that these T loop substitutions increase the flexibility of Met-$tRNA_i$ by disrupting hydrogen bonding between the T loop and the D loop, facilitating initiation at non-AUG codons (111, 203, 204). Substitutions of C3:G70, decrease initiation accuracy, increasing

UUG initiation, and disrupt TC loading, indicating that this highly conserved acceptor stem base pair helps to stabilize the open/$P_{OUT}$ conformation of the PIC, possibly in concordance with eIF1A (203). Thus, different regions of Met-$tRNA_i$ appear to function to ensure accurate translation initiation through different mechanisms.

In addition to Met-$tRNA_i$ and the 40S subunit, initiation factors also play pivotal roles in start codon recognition. These interactions have been identified largely through yeast genetic screens for $Gcd^-$ phenotypes, as well as $Sui^-$ (suppressor of initiation codon) and $Ssu^-$ (suppressor of $Sui^-$) phenotypes which affect the fidelity of start codon selection. $Gcd^-$ mutations result in constitutively derepressed *GCN4* translation, mimicking the phosphorylation of eIF2α (Fig. 1.7.5). Many classical $Gcd^-$ mutations affect eIF2 recycling or impair TC assembly, but this phenotype is also seen in mutants that impair TC loading, making it a useful tool to study factors that affect the stability of the open state to which TC binds. $Sui^-$ mutations relax the stringency of the requirement for an AUG start codon, increasing initiation at an in-frame near cognate UUG codon in a mutant *HIS4* allele lacking its AUG start codon (suppressing the $His^-$ phenotype) or in *HIS4-lacZ* reporters (increasing expression of a reporter with a UUG start codon relative to a matched AUG reporter) (Fig. 1.7.6). Such mutants generally promote transition of the PIC to its closed, arrested state, resulting in elevated initiation at suboptimal start codons. $Sui^-$ mutations often coincide with a defect in TC binding ($Gcd^-$ phenotype) because they destabilize the open state of the PIC onto which TC loads (Fig. 1.7.4). Conversely, $Ssu^-$ mutations increase the stringency of start codon recognition, preventing initiation at near cognate UUG codons, even in the presence of a $Sui^-$ mutation in another factor (suppressing the $His^+$ phenotype and decreasing expression of a *HIS4-lacZ* reporter with a UUG start codon) (Fig. 1.7.6). $Ssu^-$ mutations

generally stabilize the open conformation of the PIC, promoting continued scanning. Sui$^-$ and Ssu$^-$ mutations have been found in initiation factors 1, 1A, 2, 3, and 5, as well as Met-tRNA$_i$ and rRNA (as already mentioned above), and genetic and biochemical analyses of these mutants have shed considerable light on the mechanics of start codon selection by the scanning PIC (194). The contributions of each of these initiation factors is discussed in further detail below.

### 1.4.6 eIF1

eIF1 is a small, ~12 kD protein that binds to the platform of the 40S subunit near the P site and mRNA channel (114, 138). It is composed of an α/β core resembling several ribosomal proteins and RNA binding domains, and an unstructured NTT (205). The factor binds directly to eIF2β, the eIF5-CTD, and eIF3c, consistent with its involvement in the MFC (139). eIF1 plays a dual role in translation initiation, promoting the open conformation of the PIC that facilitates TC loading and scanning while preventing rearrangement to the closed conformation, imposing a strict requirement for AUG recognition (206). Its binding to PICs, along with eIF1A, is associated with opening of the latch at the mRNA entry channel and upward movement of the 40S head (110, 111, 127). In addition to visualizing the open latch in a cryo-EM structure of 40S bound to eIF1 and eIF1A, Passmore, et al. provide biochemical evidence that both of these factors accelerate the TC binding to the open PIC, while only eIF1A stabilizes the interaction of TC with 40S. In fact, TC was bound more tightly in the absence of eIF1, suggesting that eIF1 stabilizes the P$_{OUT}$ state while opposing transition to the P$_{IN}$ state (127). During scanning, eIF1 release is required for start codon recognition, placing eIF1 as a gatekeeper for the transition to the closed state of the PIC and start codon selection (129, 137, 207).

43S PICs require eIF1 in order to discriminate against start codons that are noncognate, are present in poor sequence context, or are <5 nucleotides from the 5' end of mRNA, and low eIF1 levels promote initiation at near cognate start codons and AUG codons in poor context genome wide (151, 208). Accordingly, mutations were identified in the yeast *SUI1* gene, encoding eIF1, that lower initiation fidelity and confer a Sui$^-$ phenotype (209). eIF1 dissociation from the PIC triggers and, along with eIF5, controls the rate of $P_i$ release from eIF2 (207, 210, 211). In Sui$^-$ mutants, in which UUG initiation is elevated, overexpression of eIF1 consistently decreases this near cognate initiation, while lowered eIF1 levels decrease initiation accuracy (12, 212). Mutations in eIF1 Loop 1 that weaken the binding of eIF1 to the 40S subunit reduce the rate of TC loading (Gcd$^-$) and elevate initiation at near cognate codons (Sui$^-$) and were shown to destabilize the open state of the PIC *in vitro* (212, 213). In contrast, eIF1 mutations that improve binding to 40S display Ssu$^-$ phenotypes by impeding transition of the PIC to the closed state (212, 214).

Recent structural evidence indicates that eIF1 bound to the open PIC would clash with the closed-state Met-tRNA$_i$ in two locations: eIF1 β-hairpin Loop 1 with the ASL and eIF1 Loop 2 with the D loop of tRNA$_i$ (91, 110, 195). Comparing the py48S-open structure with py48S-closed revealed that Met-tRNA$_i$ moves relative to eIF1 during the downward movement of the 40S head upon start codon recognition, and that the resulting predicted clashes are resolved by remodeling of eIF1 Loops 1 and 2 (111). Remodeling of Loop 1 disrupts contacts between eIF1 and the 40S and is therefore expected to weaken eIF1 binding, consistent with the Sui$^-$ and Gcd$^-$ phenotypes of Loop 1 mutants described above (212-214). Substitutions in eIF1 Loop 2 conferred Sui$^-$ phenotypes without affecting eIF1 affinity for the 40S subunit, indicating that a

clash between eIF1 and the D-loop inhibits transition of Met-tRNA$_i$ to its closed state position at non-AUG codons (200). Mutations in Loop 2 do not confer Gcd$^-$ phenotypes. Because they alleviate a clash between eIF1 and the D-loop without affecting the stability of eIF1 binding, they remove an impediment to the closed state, allowing open-to-closed transition at non-AUG codons (Sui$^-$), but they do not destabilize the open complex, and so do not affect the rate of TC loading (not Gcd$^-$).

The start codon of the eIF1 gene itself exists in poor sequence context, conserved in diverse eukaryotes, which is used as a mechanism to autoregulate cellular eIF1 levels (215). When eIF1 levels are high, the binding equilibrium shifts such that more PICs retain eIF1, increasing initiation stringency, with the result that the poor-context eIF1 AUG is skipped over with greater frequency. However, when eIF1 levels drop, eIF1 occupation on PICs decreases, allowing more frequent initiation at the eIF1 CDS. Accordingly, mutations that weaken eIF1 binding to the 40S raise eIF1 levels by relaxing the requirement for good context, while substitutions that enhance eIF1 binding increase discrimination against the eIF1 start codon, lowering eIF1 levels (212, 214). For this reason, eIF1 levels can be used to assay the effects of mutations on context recognition.

### 1.4.7 eIF1A

eIF1A is universally conserved in all kingdoms of life (216). Structurally, it shares an oligonucleotide/oligosaccharide binding fold (OB) domain with its bacterial counterpart, IF1, as well as a similar binding site in the A site of the SSU (111, 138, 217-220). In eukaryotes, this

17.4 kD protein also contains a C-terminal subdomain with two α-helices and intrinsically disordered C- terminal tail as well as an unstructured NTT (220). The CTT contains short repeats called scanning enhancer (SE) elements, which promote the open complex and accelerate TC loading in cooperation with eIF1 and prevent transition to the closed state at non-AUG codons (221). Substitutions in these repeats impart $Gcd^-$ and $Sui^-$ phenotypes *in vivo* and reduce the rate of TC loading *in vitro* (196). Evidence from hydroxyl radical mapping indicates that the eIF1A-CTT reaches into the P site and, like eIF1, might clash with Met-$tRNA_i$ in the closed PIC (220). Results from further DHRC mapping, as well as FRET assays with fluorescently labeled eIF1A and eIF1, are consistent with the notion that the eIF1A CTT is displaced from the P site following AUG recognition (137, 199). Start codon recognition triggers a strong interaction between the eIF1A CTT and the eIF5-NTD, helping to stabilize the closed conformation of the PIC (222, 223). Using FRET, Nanda et al. demonstrated that the eIF1A CTT moves toward the NTD of eIF5 upon AUG recognition, and that this movement is controlled by the rate of eIF1 dissociation from the PIC (211). An interaction between eIF1A-CTT and the CTD of eIF5 in the open, scanning PIC was indicated by NMR evidence, suggesting that the eIF1A-CTT may promote the open state while bound to the eIF5 CTD and shift to vacate the P site and interact with the eIF5 NTD following AUG recognition, allowing it to stabilize the closed complex (206, 224). SE mutants were found to uncouple eIF1 release from $P_i$ release, impairing $P_i$ release and eIF1A CTT relocation without affecting release of eIF1. These results suggest that movement of the eIF1A CTT towards the eIF5-NTD is required for $P_i$ release (211).

In addition to the scanning enhancer elements, two scanning inhibitor (SI) elements were identified in eIF1A: one in the NTT and one in the helix domain adjacent to the CTT.

Substitutions in these regions confer Ssu$^-$ phenotypes, increasing the fidelity of translation initiation (196, 223). This indicates that the SI elements promote the closed complex. Accordingly, mutation of NTT residues 17 – 21 in $SI_1$ decreased initiation at the poor context AUG of eIF1, decreasing eIF1 levels (212). All but four residues of the eIF1A NTT are visible in a recent cryo-EM structure of a partial yeast 48S PIC (py48S), revealing that basic NTT residues (K7, K10, R13, and K16) contact the mRNA directly adjacent to the AUG codon (+4 to +6) and R14 contacts 18S rRNA, placing the end of the NTT at or near the codon:anticodon duplex (110). This grants molecular insight into the mechanism of $SI_1$, which spans residues 1 – 26 of the NTT (196). Substitution of these basic NTT residues imparted Ssu$^-$ phenotypes *in vivo*, resulted in genome-wide discrimination against AUGs in suboptimal context (as measured by ribosome profiling), and destabilized PICs at UUG codons *in vitro*, suggesting that the eIF1A NTT helps to stabilize the closed conformation, and acts to promote appropriate usage of suboptimal start codons (206).

### 1.4.8 eIF5

The 405-residue, 45 kD eIF5 performs a number of functions throughout translation initiation, and its complete role in start codon selection is still being uncovered. It is composed of functionally distinct N- and C-terminal domains connected by a long, flexible linker. The NTD resembles eIF1 in structure and houses a zinc finger, while the CTD forms an α-helical HEAT domain similar to that in eIF2Bε (225, 226). The eIF5-CTD plays an important role in formation of the MFC through direct interactions with eIF1, eIF2β, and eIF3c (139, 144, 146, 147, 227). Additionally, the CTD binds the RNA2 domain of eIF4G, suggesting that it plays a role in

recruitment of 43S PICs to mRNA, possibly in cooperation with eIF3 (144, 228, 229). The interaction of eIF5 with its binding site in eIF3c is additionally proposed to stabilize the scanning PIC by preventing an interaction of this region of eIF3c with eIF1. This interaction is thought to prevent eIF1 from binding to 40S (230). The eIF5 CTD also plays an important role in start codon recognition through its interaction with eIF1. This region of eIF5 has been shown to enhance eIF1 dissociation, promoting transition to the closed PIC (210, 211, 222). NMR data demonstrates that the CTD binds to eIF2β as well as eIF1 (143). Mutation of the eIF5-CTD:eIF2β interface in a manner that does not disrupt eIF5-CTD binding to eIF1 results in inhibition of eIF1 release from the PIC and impaired start codon recognition, suggesting that eIF5-CTD:eIF2β interaction promotes the closed state by prompting eIF1 dissociation (143). More recently, it was demonstrated that the CTD binds to eIF2 and TC with affinity equal to the full eIF5 and promotes eIF1 release to the same extent seen for full-length eIF5, while the eIF5 NTD has no effect (211). In addition, overexpression of eIF5 causes increased eIF1 release, resulting in a Sui$^-$ phenotype in yeast (210) and mammals (231). Like eIF1, eIF5 is autoregulated by a translational mechanism in which high eIF5 levels induce initiation at the poor-context AUG of an inhibitory uORF in eIF5 mRNA (231). Thus, it appears that the eIF5-CTD facilitates eIF1 release and stabilizes the closed PIC preferentially at AUG codons in appropriate sequence context. Thus, broadly speaking, eIF5 and eIF1 have opposing functions in controlling initiation accuracy, with eIF1 discriminating against, and eIF5 promoting, usage of poor initiation codons.

The eIF5 NTD houses the GTPase activating protein (GAP) domain that stimulates hydrolysis of GTP on eIF2 and has been shown to bind directly to the eIF2γ G domain (232, 233). This GAP activity depends on eIF5-R15 in the unstructured N-terminus, mutation of which is lethal and

imparts a dominant defect in GTP hydrolysis (121, 207, 234). eIF5-R15 is predicted to interact with the GTP binding pocket on eIF2γ to allow GTP hydrolysis by stabilizing the necessary intermediate (112). As noted above, a combination of FRET and kinetic studies revealed that movement of the eIF1A CTT and eIF5 NTD toward each other is required for release of both $P_i$ and eIF1, and is dependent on SE elements in the eIF1A CTT (211, 222). Mutations in the SE elements decoupled $P_i$ release from eIF1 dissociation (211). Consistent with this, the lethal, dominant $Sui^-$ mutant eIF5-G31R has been shown to strengthen the binding of eIF1A to PICs at near-cognate UUG codons, and this mutant also alters $P_i$ release and stabilizes closed state PICs (214, 222, 235, 236).

Until recently, there was little structural information available on the position of eIF5 within the PIC. Based on genetic and biochemical data, it was previously suggested that release of eIF1 might prompt movement of eIF5 to the vacated eIF1 binding site, and that this movement might allow it to stimulate $P_i$ release (210). A recent cryo-EM structure reveals a PIC lacking eIF1 in which the eIF5-NTD is bound near the P site in the eIF1 binding site, supporting this hypothesis (112). Thus, it appears that eIF5 may impart irreversibility to start codon selection by occupying the binding site of eIF1 following AUG recognition and preventing that factor's rebinding. In this position, the NTD interacts extensively with Met-$tRNA_i$, stabilizing it in a conformation thought to be an intermediate between start codon recognition and subunit joining. In this state, Met-$tRNA_i$ is bound deeply in the P site, but tilted toward the 40S body compared to previous $P_{IN}$ structures, presumably in preparation for eIF5B binding. Mutations predicted to disrupt interactions between Met-$tRNA_i$ and the eIF5-NTD based on this structure conferred $Ssu^-$ phenotypes *in vivo*, and disfavored transition to the closed PIC and also reduced the rate of $P_i$

release *in vitro*, all in the presence of a Sui$^-$ variant of eIF2β (encoded by *SUI3-2*), supporting the role of eIF5-NTD in stabilizing the closed state (112). Similarly, eIF5-NTD mutations expected to stabilize interactions of the eIF5-NTD with the PIC result in Sui$^-$ phenotypes, most notably in the case of eIF5-G31R, which confers a strong Sui$^-$ phenotype by preferentially accelerating $P_i$ release and stabilizing the closed state at UUG codons, potentially by preventing interaction of the eIF5-NTD with the eIF1A-CTT.

In addition to its roles in PIC assembly and start codon recognition, eIF5 acts as a GDP dissociation inhibitor (GDI) by tightly binding eIF2•GDP following $P_i$ release, preventing any spontaneous release of GDP and ensuring recycling through eIF2B (237). This activity requires the eIF5-CTD and central linker region, but not the NTD. As mentioned above, the CTD and linker interact with eIF2γ and eIF2β (233, 238, 239). eIF5 and eIF2B share a conserved eIF2 binding site, and approximately 40% of cellular eIF2 is sequestered in complex with eIF5 (240, 241). Overexpression of eIF5 impedes TC formation, an effect intensified by mutations in eIF2B, whereas overexpression of the catalytic subunit of eIF2B exacerbates eIF5 mutations, indicating competition between eIF5 and eIF2B for binding eIF2 (240). It is proposed that eIF5, in its role as GDI, dissociates from the PIC in a complex with eIF2•GDP and delivers eIF2•GDP to eIF2B, which acts as a GDI dissociation factor (GDF) to release eIF5 from eIF2•GDP and permit eIF2 cycling by eIF2B (238, 242, 243). The GDI function of eIF5 makes a critical regulatory contribution to the inhibition of GEF activity by eIF2α phosphorylation (237). Indeed, overexpression of eIF5 mimics eIF2α phosphorylation, activating *ATF4* translation in mammals (244). While this activation could be the result of decreased TC availability resulting from impaired eIF2 recycling, as described above, it is also possible that eIF5 plays a more direct role

in the important translational control pathway mediated by eIF2α phosphorylation. This system is made more complex in mammalian cells and other higher eukaryotes by the presence of an eIF5 mimic protein (5MP), whose competition with eIF5 for eIF2 binding has been proposed to affect the rate of initiation at non-AUG start codons (245).

### 1.4.9 eIF3

eIF3 is the largest and most complex initiation factor, consisting of five subunits in yeast and 13 in most multicellular eukaryotes (229, 246-249). This multisubunit complex has been implicated in nearly every step of translation initiation (135, 250). In *S. cerevisiae*, eIF3 is composed of five core subunits (eIF3a, eIF3b, eIF3c, eIF3g, and eIF3i), which are required for translation *in vivo* (229, 250-252). A sixth subunit, eIF3j, is nonessential and is present at substoichiometric levels. Although research into the functions of eIF3j is ongoing, it has been shown to enhance interaction of eIF3 with other initiation factors and promote eIF3 binding to the 40S subunit, as well as performing a role in processing 20S pre-rRNA (175, 253-256).

While the core of eIF3 binds on the solvent face of the 40S subunit, extensions from the eIF3a-CTD and eIF3c-NTD nearly encircle the SSU, reaching from the mRNA entry channel around the solvent face past the mRNA exit channel to near the P site (109, 111, 114, 116, 249). This expansive positioning helps to explain the diverse roles of eIF3, from involvement in TC assembly and 43S PIC formation to scanning and stabilizing the 48S PIC (137, 139, 145, 189, 255, 257, 258). It has been suggested that eIF3 may act to coordinate these events (135). Genetic and biochemical data implicate eIF3 in interactions with eIF1, eIF1A, and eIF5, as well as with

the TC (128, 139, 145, 146, 205, 229). It has also been demonstrated that eIF3 stays bound to ribosomes during the translation of short uORFs and facilitates reinitiation at downstream ORFs (259). Recent data further implicate eIF3 in translation termination (135). Much of the structure is flexible, and while recent structural studies have revealed considerable new information about the shape and position of eIF3, several regions remain unresolved. Research continues into the many facets of this far-reaching factor.

eIF3 interacts with 40S subunits near both the mRNA entry and exit channels (109, 260-263). It has been shown to crosslink to mRNA at nucleotides -8 to -17 in the exit channel (numbered from the A at +1 of an AUG positioned in the P site) (261). In addition, a mutation in the NTD of eIF3a, a segment located near the exit channel, perturbs mRNA recruitment to the 43S PIC in yeast (258). Recent high resolution cryo-EM structures of eIF3 bound to a partial yeast 43S or 48S PIC reveal a heterodimer of PCI (Protease/Cop9/eIF3) domains comprised of the eIF3a NTD and eIF3c CTD bound near the mRNA exit channel, while eIF3i, eIF3g, the eIF3a CTD, and the C-terminal region of eIF3b form a separate subdomain bound near the entry channel (111, 114, 136). An analysis of mRNA recruitment in an *in vitro* system using mRNAs designed to leave either the entry or exit channel empty revealed that contacts near the entry channel play a role in TC binding to the PIC and facilitating mRNA recruitment, while contacts near the exit channel stabilize the binding of mRNA in the 48S PIC (136). Recent studies from our lab demonstrated that basic residues of the entry channel ribosomal protein Rps3 in proximity to the mRNA are functionally redundant with eIF3 in stabilizing mRNA contacts with the PIC, indicating that eIF3 may also function at the entry channel to stabilize mRNA binding to the PIC in concert with ribosomal proteins (119).

### 1.4.10 Subunit joining

Following AUG recognition, the 48S PIC in the closed conformation contains Met-tRNA$_i$ base paired to AUG in the P site and retains the initiation factors bound before start codon selection, except for eIF1. Based on structural data, it is apparent that eIF2•GDP must dissociate from the intersubunit face prior to subunit joining. eIF2•GDP dissociation is likely aided by its reduced affinity for Met-tRNA$_i$ following $P_i$ release, and there is evidence that it dissociates in complex with eIF5, as described above (123, 237). The timing of eIF3 release is not known, although there is evidence, noted above, that it remains bound to the translating ribosome long enough to induce reinitiation following translation of short ORFs (264). Rather than dissociate, eIF1A becomes more tightly bound to the 48S PIC following eIF1 release (110), and stimulates the final step of initiation, joining of the 60S subunit, as described next.

eIF2•GDP dissociation is stimulated by binding of the GTPase eIF5B, which promotes 60S subunit joining (265-267). eIF5B is a universally conserved initiation factor, the homolog of IF2 in bacteria, and although it is not essential, deletion of its structural gene *FUN12* in yeast results in a strong translation initiation defect (266, 268). Although the 40S and 60S subunits naturally bind each other *in vitro*, particularly at high concentrations of magnesium ions, the presence of other initiation factors in the 48S PIC necessitate eIF5B to stimulate this reaction, both *in vivo* and *in vitro* (142, 267). The crystal structure of the four-domain archaeal aIF5B exhibits a cup-like structure consisting of a G domain, domain II, and domain III, while domain IV forms a kind of base, connected to the cup by a long α-helix (194, 269, 270). Compared to archaea, both eIF5B and the bacterial IF2 contain an additional N-domain which is poorly conserved and was

shown not to be necessary for eIF5B function (271). Interaction of the unstructured eIF1A CTT assists the binding of eIF5B to the ribosome through interaction with the eIF5B CTD, in an interaction unique to eukaryotes, which stimulates subunit joining as well as hydrolysis of eIF5B-bound GTP (128, 272-274). Crystallographic and NMR data confirm an interaction of the eIF1A-CTT with domain IV of eIF5B (275, 276).

Binding of GTP results in conformational changes to eIF5B, that may include release of domain III, allowing greater flexibility throughout domains III and IV (270). Domain IV may then be stabilized by binding to Met-tRNA$_i$, explaining the lack of catalysis by eIF5B in the absence of Met-tRNA$_i$ in the P site (274). This interaction appears to depend on the methionylated 3'-CCA end of Met-tRNA$_i$ (115). Binding of eIF5B may stabilize Met-tRNA$_i$ on 48S PICs following eIF2•GDP release, preventing either Met-tRNA$_i$ dissociation or continued scanning of the PIC (194). Consistent with this, yeast genetic experiments revealed that 48S PICs are less stable and that leaky scanning is increased *in vivo* in the absence of eIF5B (277). GTP hydrolysis is required for release of eIF5B and eIF1A from 80S PICs following subunit joining, and it is believed to alter the 80S conformation to allow eIF1A release (274). Mutations that impair GTP hydrolysis by eIF5B prevent eIF1A dissociation and result in severe growth phenotypes in yeast, although they do not prevent subunit joining (271, 278-280). After the release of these factors, the 80S ribosome is poised to begin elongation.

Together, these data present the following detailed model of start codon selection by the scanning mechanism in eukaryotes. eIF1 and eIF1A bind to 40S subunits, possibly as members of a multi-factor complex with eIF3 and eIF5, and promote the open conformation of the SSU

(Fig. 1.7.7 (i)). eIF2-GTP in ternary complex with Met-tRNA$_i$ is recruited to the PIC by interactions with members of the MFC. This open conformation, together with the helicase activity of eIF4A, recruited to the mRNA cap as a subunit of eIF4F complex, eIF4B, and interactions of eIF4G with 40S-bound eIF3 or eIF5, enable recruitment of the newly formed 43S PIC to the 5' end of mRNA. The 43S PIC then scans along the mRNA toward the 3' end, assisted by the helicase activity of eIF4A, enhanced by eIF4B, and by helicase Ded1. The GTP on eIF2 may be hydrolyzed during scanning, but the inorganic phosphate is not released (Fig. 1.7.7 (ii)). When an AUG in good sequence context is encountered, it base pairs to the anticodon of Met-tRNA$_i$, causing a conformational shift of the tRNA, which is bound deeper into the P site ($P_{IN}$). This movement results in clashes of Met-tRNA$_i$ with eIF1 Loops 1 and 2 and with the eIF1A-CTT (Fig. 1.7.7 (iii)). These clashes are likely alleviated by deformation of eIF1 Loops 1 and 2 and displacement of eIF1 on the 40S platform (Fig. 1.7.7 (iv)), leading to eIF1 dissociation from the 40S subunit. Release of eIF1 triggers movement of the eIF1A CTT from the P site and interaction with the eIF5 NTD. Dissociation of the eIF1A-CTT from the eIF5-CTD enables the latter to interact with eIF2β, and the alternative interaction of eIF1A-CTT with the eIF5-NTD gates $P_i$ release, allowing completion of GTP hydrolysis on eIF2. At the same time, the NTD of eIF5 shifts to occupy the eIF1 binding site, preventing eIF1 reassociation (Fig. 1.7.7 (v)). The eIF2•GDP dissociates from Met-tRNA$_i$ and is released from the PIC in a complex with eIF5. eIF5B then binds the 48S PIC, aided by interaction with the eIF1A CTT, and possibly through contacts with Met-tRNA$_i$, and facilitates joining of the 60S subunit. Hydrolysis of the GTP bound on eIF5B allows release of eIF5B and eIF1A. Thus, the end result of translation initiation is an 80S ribosome containing a Met-tRNA$_i$ base paired to a start codon in the P site, possibly bound by eIF3, and ready to continue into translation elongation. In spite of the many molecular

interactions already determined to play a role in start codon recognition, much work remains to determine the order of events and the energetics of the process, and to uncover additional contacts involved in the network of interactions that determine the accuracy of start codon selection. In this work, we examine the roles of eIF2β and Rps26/eS26 in this process, and determine that they work to stabilize the open and the closed states of the PIC, respectively, expanding our understanding of the nuanced mechanism that governs the recognition of an appropriate start codon.

**1.5 eIF2**

An important mechanism for regulating both global and mRNA-specific translation initiation involves phosphorylation of the α-subunit of eIF2. The eIF2 is a heterotrimeric G protein made up of α, β, and γ subunits, encoded in yeast by the genes *SUI2*, *SUI3*, and *GCD11*, respectively (281-283). In its active, GTP-bound state, eIF2 binds Met-tRNA$_i$ and delivers it to the 40S ribosome in a ternary complex (TC), as described above. Through phosphorylation of the α subunit, which reduces its activity, eIF2 is also a key regulator of translation initiation. Although there is no bacterial ortholog of eIF2, it is well conserved throughout archaea and eukaryotes. Recent cryo-EM structures of the PIC, as well as an archaeal crystal structure of aIF2, reveal that eIF2γ forms the core of the complex to which the α and β subunits bind (Fig. 1.7.8 A) (110, 111, 284, 285). There is no evidence of direct interaction of eIF2α and eIF2β in yeast or archaea (110, 111, 285, 286). Possible interactions between these two subunits have been reported in a reconstituted human system (287).

eIF2α is made up of an N-terminal oligonucleotide/oligosaccharide binding (OB)-fold domain that interacts with the Met-tRNA$_i$ ASL and contains the conserved phosphorylation site, serine 51 (S51), in a mobile loop, a central α-helical domain that contacts the D- and T-loops of tRNA$_i$, and a C-terminal α/β domain, which contacts the tRNA$_i$ acceptor stem and binds eIF2γ, and is connected to the α-helical domain via a long, flexible linker (110, 111, 285, 288, 289). The features of eIF2β that are well conserved among archaea and eukaryotes include an N-terminal α helix that binds tightly to eIF2γ, a central helix-turn-helix (HTH) domain, and a C-terminal zinc binding domain (ZBD) (286, 290-292). The eukaryotic protein also includes a presumably unstructured N-terminal extension that contains three lysine-rich K-box elements, as well as a short (~15 residue) C-terminal extension (293). Before this work, much of the structure of eIF2β remained unresolved. At 57 kD, eIF2γ is the largest of the subunits and acts as the core of the eIF2 complex. It consists of an N-terminal domain housing the GTP binding site and two β-barrel domains in both archaea and eukaryotes, with a eukaryote-specific N-terminal extension of up to 90 residues (194, 293). The G domain of eIF2γ is similar in amino acid sequence and structure to the bacterial elongation factor EF-Tu, which facilitates the binding of aminoacyl-tRNAs in the ribosomal A site (283, 294). eIF2γ serves a similar function, although with greater specificity, by binding Met-tRNA$_i$ in the P site during initiation (194).

### 1.5.1 eIF2 recycling by eIF2B

eIF2 is recycled in successive rounds of translation by replacement of GDP on eIF2•GDP with GTP (Fig. 1.7.5 A). The GTP-bound form of eIF2 has a 50- to 100-fold greater affinity for Met-tRNA$_i$ than the GDP-bound form (123, 207). eIF2•GTP thus forms a ternary complex with Met-

$tRNA_i$ and delivers it to the ribosome during translation initiation. Hydrolysis of eIF2-bound GTP occurs during scanning, requiring eIF5 as a GTPase activating protein (GAP), and is completed with $P_i$ release following start codon recognition. The resulting eIF2•GDP leaves the ribosome with eIF5, which also acts as a GDP dissociation inhibitor (GDI) to prevent unregulated release of GDP from eIF2. eIF5 delivers eIF2•GDP to its guanine nucleotide exchange factor (GEF), eIF2B, which acts as a GDI displacement factor (GDF), displacing eIF5, all as described above. eIF2B consists of five subunits, α – ε, with the GEF domain housed in the ~200 C-terminal residues of the ε subunit. Cellular levels of GTP are much higher than those of GDP. eIF2B catalyzes guanine nucleotide exchange by decreasing the binding affinity of eIF2 for GDP, allowing eIF2 to bind the readily available GTP and begin a new cycle of initiation by binding Met-$tRNA_i$ (reviewed in (194)).

### 1.5.2 Integrated Stress Response

In addition to its role in various disease states, the regulation of translation also plays a critical role in the cellular response to a variety of stressors, including temperature stress, nutrient stress, ER stress, and viral infection. Surprisingly, these disparate stressors were found to down-regulate translation through a central mechanism known as the integrated stress response (ISR), mediated by the key mechanism for translational control involving phosphorylation of eIF2α on S51 (reviewed in (295, 296)). In mammals, there are four eIF2α kinases, each activated by a distinct stressor. PERK responds to ER stress resulting from unfolded proteins, HRI is activated by heme deprivation, PKR mediates the cellular response to the presence of double stranded RNA during viral infection, and GCN2, the only eIF2α kinase conserved in yeast, is activated by nutrient

stress. These four kinases, each of which phosphorylates the same residue (S51) of eIF2α, constitute the Integrated Stress Response (ISR), reducing bulk translation while upregulating specific stress response genes. Phosphorylation of eIF2α leads to an excessively stable interaction with eIF2B, sequestering the latter in an inactive complex and preventing recycling of the unphosphorylated eIF2•GDP to eIF2•GTP. The ensuing reduced concentration of eIF2•GTP and attendant decrease in TC formation results in a global down-regulation of translation. This decrease in protein synthesis is an appropriate response to many stressors, conserving resources and preventing the translation of viral proteins. However, the expression of certain stress response genes increases during ISR. This apparent paradox can be explained by the unique translational control mechanism of the gene *GCN4* (*ATF4* in mammals).

*GCN4* is a transcription factor that transcriptionally activates amino acid biosynthetic genes. The *GCN4* mRNA contains four upstream open reading frames (uORFs) 5' of the main CDS. The first short uORF from the mRNA 5' end (uORF1) encodes a three amino acid product and contains sequences that promote continued ribosomal scanning rather than recycling following termination at uORF1. After translating uORF1, the post-termination 40S ribosomal subunit resumes scanning and must acquire a new TC before it can reinitiate translation. Under normal cellular conditions, TC is readily available and loads rapidly onto the scanning ribosome, allowing it to reinitiate at uORF2, 3, or 4. After translating any of these uORFs, termination of translation will result in recycling of both ribosomal subunits from the mRNA. As a result, very few ribosomal complexes are able to scan to the *GCN4* coding sequence, and *GCN4* expression is extremely low (Fig. 1.7.5 B (i)). However, when eIF2α is phosphorylated in response to stress conditions, TC levels are reduced, increasing the likelihood that, following translation of uORF1,

the re-scanning ribosome will scan past all 3 remaining *GCN4* uORFs before acquiring TC. In this case, the ribosome will reach the *GCN4* start codon and, if TC has been loaded, it will initiate and synthesize the *GCN4* gene product (Fig. 1.7.5 B (ii)). Through this unique process of translational control, the *GCN4* gene is upregulated under stress conditions which globally downregulate translation. The GCN4 protein then promotes the transcription of other genes needed by the cell to mount a stress response. In mammals, *ATF4* activation also results in the downstream activation of GADD34, an eIF2α phosphatase, providing negative feedback.

Gcd$^-$ mutations in yeast activate the ISR in the absence of cellular stressors by constitutively derepressing *GCN4*. This phenotype is conferred by any mutation that, like eIF2α phosphorylation, allows ribosomes to scan to the *GCN4* coding sequence following translation of uORF1 before obtaining a new TC, whether by decreasing TC abundance or by impairing the binding of TC to the PIC (Fig. 1.7.5 B (iii)). Gcd$^-$ mutations have been obtained in all of the subunits of eIF2B, which likely impair eIF2 recycling and mimic eIF2α phosphorylation, in the three subunits of eIF2 and Met-tRNA$_i$, which likely reduce TC assembly, and in eIF1, eIF1A, and eIF3 subunits that likely reduce the on-rate of TC binding to 40S subunits.

### 1.5.3 eIF2β

The β subunit of eIF2 is involved in the process of eIF2 recycling and the ISR by anchoring eIF5 and eIF2B onto the eIF2 complex through interaction with its eukaryotic-specific N-terminal K boxes. Deletion or disruption of all three K boxes of eIF2β is lethal, although deletion of one or two does not affect cell viability (297, 298). Removal of the K boxes conferred an mRNA

binding defect, but did not impact binding of Met-tRNA$_i$ to eIF2 *in vitro* (298). A conserved bipartite motif of acidic and aromatic amino acids located at the C terminus of both eIF5 and eIF2Bε was shown to mediate the binding of these factors to the N-terminus of eIF2β (297). eIF2β also plays an important role in binding Met-tRNA$_i$, requiring the N- and C-terminal domains for maximal binding, with the C-terminal extension of eIF2β having the greatest impact on binding affinity (299). Additionally, residues of the N-terminal α-helix are required for binding to eIF2γ and thus eIF2 complex formation (300). The function of the C-terminal zinc binding domain (ZBD) in eIF2β has not been elucidated. Deletion of the ZBD imparts dominant Gcd$^-$ and recessive lethal phenotypes and impairs mRNA binding to eIF2β (298, 301). Point mutations of any of the cysteines of the zinc finger or the adjacent residues also result in recessive lethal phenotypes and confer dominant Sui$^-$ phenotypes (282, 301). These mutations (*SUI3-2*, encoded by S264Y, and L254P) display a higher intrinsic rate of GTP hydrolysis, independent of eIF5, which might underlie the increased initiation at near cognate UUG codons (235); however, this has been disputed. Because of the proximity of these residues to the eIF2γ G domain, it is possible that they normally act to prevent GTP hydrolysis or P$_i$ release at non-AUG codons (12).

Before the current work, little else was known about the role of eIF2β in start codon recognition. Previous structures of the eukaryotic ribosome did not include eIF2β (109, 110). New structural evidence, discussed in Chapter 2, visualized the previously unresolved eIF2β in the open and closed structure of the PIC (py48S-open and py48S-closed) (111). These structures revealed that the helix-turn-helix (HTH) domain of eIF2β contacts eIF1 and eIF1A in the open structure, while it is positioned away from these factors in the closed state (Fig. 1.7.8 B). Differential contacts

between eIF2β and Met-tRNA$_i$ in the open and closed states were also visualized for the first time (Fig. 1.7.8 C). These observations allowed us to predict residues of eIF2β, substitutions of which would alter exclusively the stabilization of either the open or closed state, as described in Chapter 3. Genetic and biochemical analyses based on these predictions revealed roles for eIF2β in anchoring eIF1 onto the open complex before AUG recognition, and in stabilizing Met-tRNA$_i$ in the open complex in the absence of a fully formed P site and tRNA$_i$ contacts with the 40S body that exist in the closed state. In addition, eIF2β in the open PIC prevents transition of Met-tRNA$_i$ into its closed state position through steric and electrostatic clashes (302). These findings indicate a model in which Met-tRNA$_i$ is held in place by eIF2β, which also aids binding of eIF1 onto the complex, promoting the open conformation. Following start codon recognition, eIF2β moves away from eIF1 and the Met-tRNA$_i$ ASL during head closure, destabilizing the open state, promoting eIF1 release, and allowing Met-tRNA$_i$ to transition to the closed state (Fig. 1.7.8 D). These roles of eIF2β add to a growing network of interactions within the PIC that restrict initiation to appropriate start codons.

## 1.6 Context Recognition

While bacteria recognize a start site through direct base pairing to the Shine-Dalgarno sequence, the scanning mechanism of eukaryotic initiation relies on recognition of the AUG start codon by the anticodon of tRNA$_i$, as described above. This process is aided by the nucleotides surrounding the start codon, referred to as context nucleotides. Optimal sequence context increases the likelihood of initiation from a given start codon, and is especially important for initiation at non-AUG codons, such as UUG, where it presumably compensates for imperfect base pairing to the

Met-tRNA$_i$ anticodon (303). Marilyn Kozak identified the consensus sequence gccRccAUGG, which is preferred in most vertebrates, where lower case letters represent more variability, and R is a purine (A or G) (Fig. 1.7.9 A) (304, 305). 97% of vertebrate mRNAs contain a purine at the -3 position (numbered from the A of AUG (+1)), and genetic analysis confirmed the importance of a purine at this position for start codon selection (306, 307). Although the consensus sequence differs in plants, ~80% of plant mRNAs contain purines at the -3 and +4 positions (308). Budding yeast also display a preference for a purine at the -3 position, usually A, and the rest of the optimal sequence context in yeast is simply poly(A) (309). Analysis of protein abundance resulting from yeast mRNAs differing in only one context nucleotide confirmed the importance of the contribution of context nucleotides in the -1 to -3 positions (310). Mutation of the 5'UTR upstream of the Kozak sequence (-9 to -15) revealed that this upstream region also impacts translation initiation (311).

Despite the importance of the Kozak sequence for initiation efficiency, analysis of eukaryotic genomes reveals many genes that are initiated in poor sequence context. This raises the possibility that sequence context plays a regulatory role, allowing these genes to be initiated less frequently or under certain conditions. Indeed, it is known that eIF5 is regulated by such a mechanism (231). It has been observed that genes with non-optimal start codons also contain more secondary structure in their 5'UTRs (312). A recent study of Drosophila genes containing strong or weak Kozak context indicated that genes with weak sequence context were more sensitive to a decrease in initiation rate, but less sensitive to changes in elongation rate compared to genes with a strong initiation sequence context, providing a possible mechanism by which genes containing a start codon in poor context could be differentially regulated (313). It has been

suggested that such specific usage of genes initiating in poor context could mediate cellular responses to stress and other conditions, as well as playing a role in health and disease. However, in spite of the important role that context nucleotides play in start codon recognition, it is not known how the sequence context is recognized by the scanning ribosome.

When an AUG is positioned in the P site, the important -1 to -3 context nucleotides lie in the E site, and the upstream context (-4 onward) is located in the exit channel (Fig. 1.7.2 C). eIF2α contacts the -3 nucleotide, potentially enforcing the requirement for a purine at this position (Fig. 1.7.9 B) (110, 265). eIF2 complexes lacking eIF2α reduced the stimulatory effect on PIC assembly of a purine at the -3 position in a reconstituted mammalian system (265). In addition, the β-hairpin of Rps5/uS7, which also contacts the -3 nucleotide, is required for efficient and accurate start codon selection (Fig. 1.7.9 B) (314). As well as its direct contacts with mRNA, Rps5 also contacts eIF2α, forming differential contacts in the open and closed PIC that stabilize each state, respectively (315, 316). Thus, eIF2α and Rps5 may work together in context recognition, which would implicate both an initiation factor and ribosomal protein in this process.

In addition to the β-hairpin of Rps5/uS7, the SSU proteins Rps14/uS11, Rps26/eS26, and Rps28/eS28 contact mRNA in the exit channel (Fig. 1.7.9 C) (91). Of these, Rps5/uS7 and Rps26/eS26 directly contact the important -3 context nucleotide (Fig. 1.7.9 B, D) (195). Crosslinking studies implicate Rps26 in binding mRNA further upstream, at nucleotides -4 to -9, and more recent structural data supports contacts between Rps26 and such upstream context nucleotides (110, 317). Recent data indicate that Rps26 is required to maintain the preference for

the optimal Kozak sequence in yeast (318). Further, depletion of Rps26 was found to activate stress response pathways for high salt and pH stress, while ribosomes lacking Rps26 were generated in response to these stressors (318). Beyond supporting a role for Rps26 in recognition of start codon context, these data indicate a potential translational response mechanism to preferentially express genes with poor Kozak context under stress conditions.

In order to examine the role of Rps26 in recognition of start codon context, we generated substitutions in highly conserved Rps26 residues implicated in crosslinking to mRNA, as well as those predicted to contact the -3, -4, and -5 nucleotides (Fig. 1.7.9 D). As described in Chapter 4, the loss of basic (K66, K70) or polar (Y68) residues contacting the mRNA upstream of the -9 nucleotide, and particularly the introduction of acidic residues at these positions, resulted in strong Ssu$^-$ phenotypes, suggesting a role for Rps26 in stabilizing the closed PIC by binding to mRNA in the exit channel. Mutation of the region of Rps26 near the E site altered start codon selection in unexpected ways. Many substitutions in this region resulted in weak Sui$^-$ phenotypes (e.g., H80E, A81K), but displayed an apparent decoupling of AUG recognition from the preference for optimal start codon context, indicating that these phenotypes may result from different mechanisms, unlike most other known Sui$^-$ mutations. Other substitutions resulted in Ssu$^-$ phenotypes (N25A/H80A, A81D). These findings indicate a nuanced role for Rps26 in start codon and context selection that will require further study to fully dissect.

Translation initiation in eukaryotes is a complicated process that depends on a complex network of interactions among rRNA, ribosomal proteins, initiation factors, and tRNA. Changes to these interactions alter the stringency of start codon selection and recognition of Kozak context, thus

altering the proteome. While the fidelity of initiation must be safeguarded, a growing body of evidence suggests that in many cases, modulation of initiation fidelity acts as a regulatory mechanism, resulting in differential protein expression. Dysregulation of this balance is associated with multiple disease states and congenital syndromes. Thus, a full understanding of the mechanisms of start codon selection is critical. Here, we demonstrate a dual role for eIF2β in start codon selection, in stabilizing the open PIC through interactions with Met-tRNA$_i$ and by binding eIF1, while also preventing the transition of Met-tRNA$_i$ to the closed state prior to AUG recognition through steric and electrostatic clashes. Further, we provide evidence that Rps26 stabilizes the closed PIC following AUG recognition through its interactions with mRNA in the exit channel, possibly in cooperation with Rps3 in the entry channel, while playing a nuanced role in the recognition of Kozak context. These findings join a growing network of interactions that control and fine-tune the process of translation initiation to maintain the exact cellular protein levels required for cell viability, stress response, or development.

## 1.7 Figures and Tables

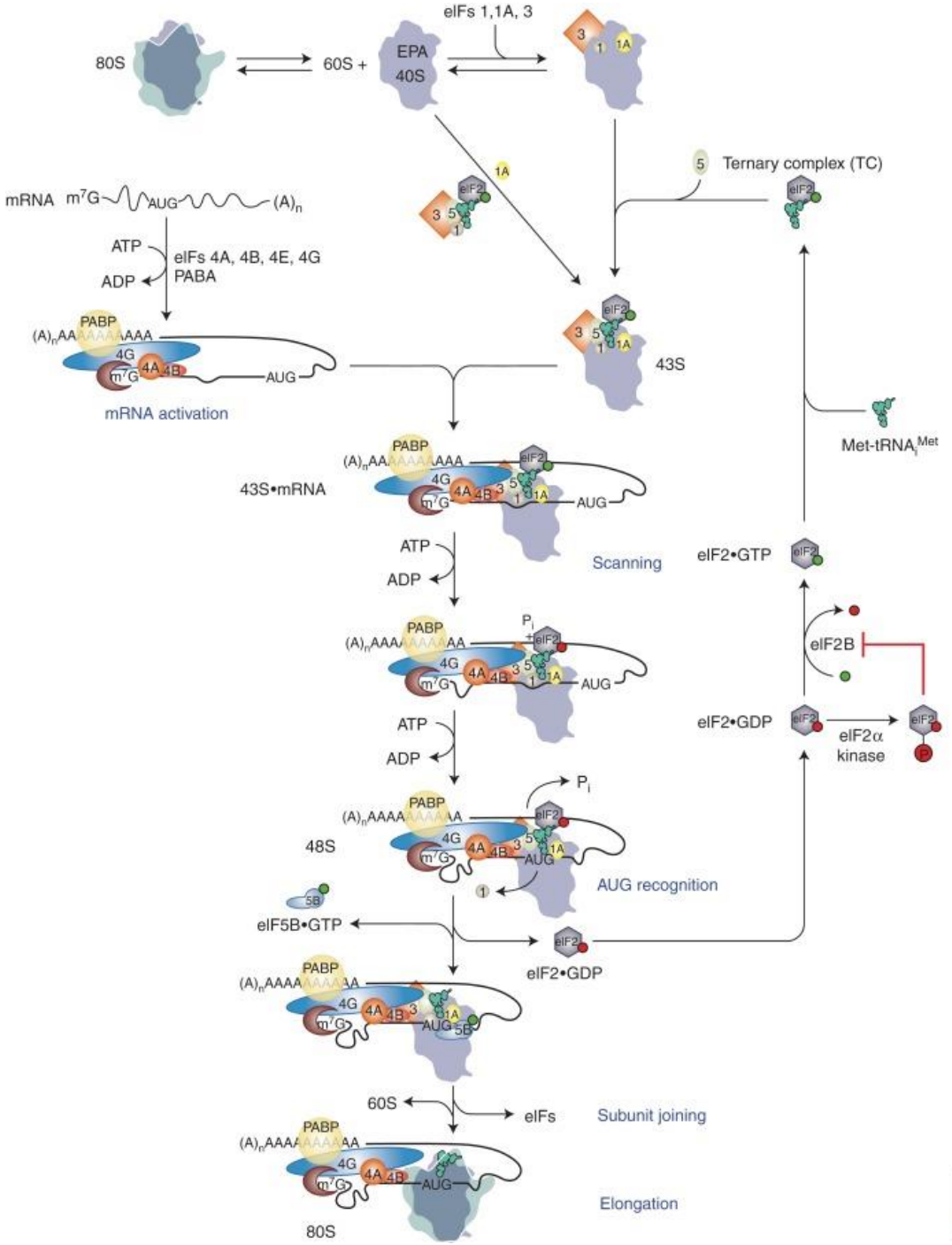

80S
60S +
EPA
40S
eIFs 1,1A, 3
Ternary complex (TC)
mRNA
ATP
ADP
eIFs 4A, 4B, 4E, 4G
PABA
PABP
AUG
mRNA activation
43S
43S•mRNA
Scanning
Met-tRNA$_i^{Met}$
eIF2•GTP
eIF2B
eIF2•GDP
eIF2α
kinase
$P_i$
48S
AUG recognition
eIF5B•GTP
eIF2•GDP
60S
eIFs
Subunit joining
Elongation
80S

**Figure 1.7.1. Model of translation initiation in eukaryotes.**

Initiation begins with assembly of a ternary complex (TC) consisting of GTP-bound eIF2 and initiator methionyl tRNA (Met-tRNA$_i$) that binds to the 40S small ribosomal subunit either on its own or as part of a multifactor complex (MFC) with eukaryotic initiation factors (eIFs) 1, 1A, 3, and 5, forming a 43S preinitiation complex (PIC). The eIF4F complex, consisting of the cap-binding protein eIF4E, scaffolding protein eIF4G, and RNA helicase eIF4A binds the 5' cap of mRNAs and forms a closed loop structure through interactions between eIF4G and the poly-(A) binding protein (PABP). The PIC is recruited to the 5' end of mRNA via interactions with the eIF4F complex and scans in a 5' to 3' direction until an appropriate start codon is recognized. Following start codon recognition, eIF1 is released, hydroplysis of the GTP on eIF2 is completed with the release of inorganic phosphate, and structural changes in the PIC prevent further scanning. eIF5B binds to the resultant 48S PIC and recruits the large 60S subunit in preparation for elongation. eIF2•GDP is released from the complex and recycled via guanine nucleotide exchange facilitated by its GEF, eIF2B. (Hinnebusch and Lorsch, 2012)

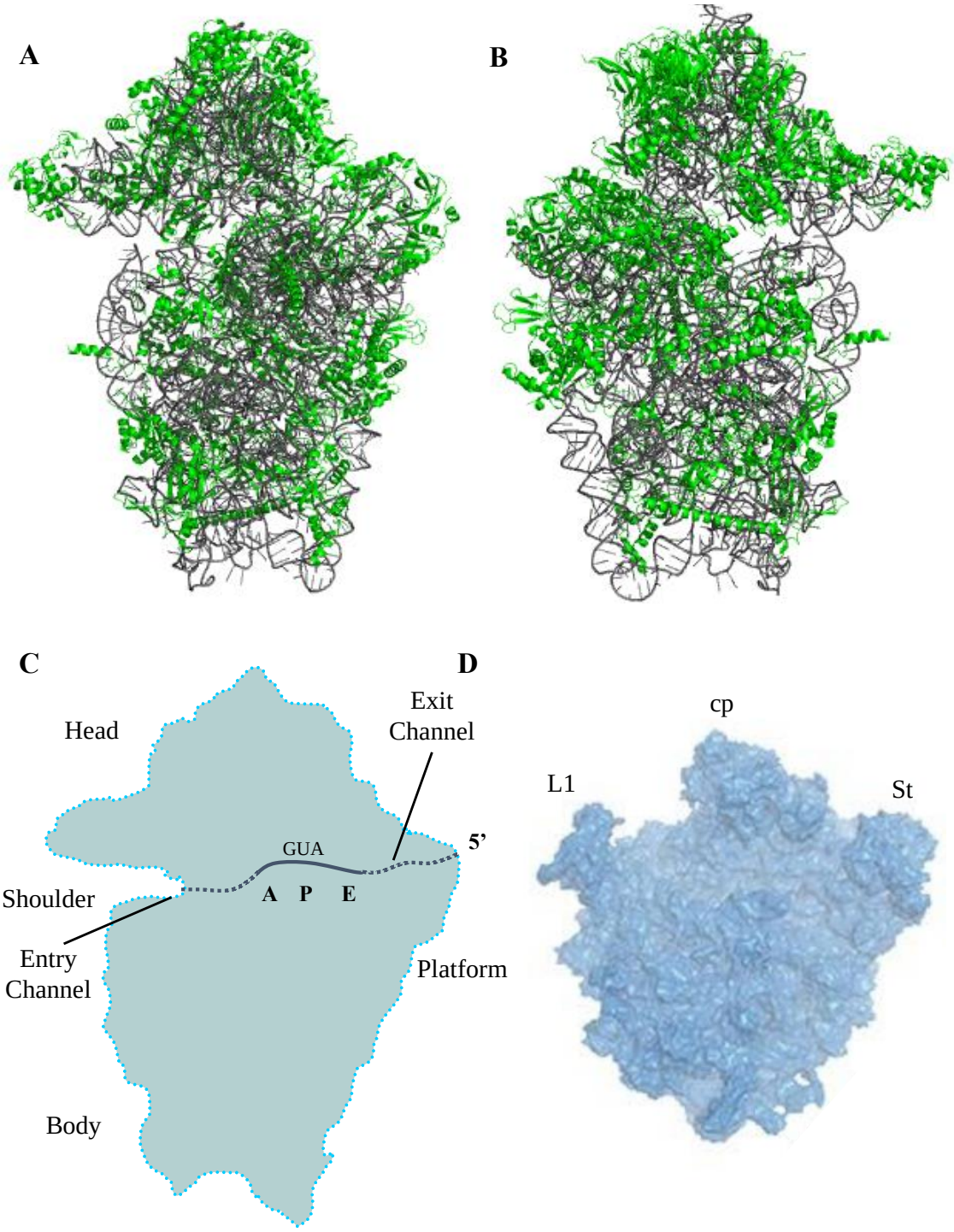

A
B
C
D
Head
Exit
Channel
5'
GUA
A P E
Shoulder
Entry
Channel
Platform
Body
cp
L1
St

**Figure 1.7.2. Ribosomal structure.**

**(A)** Solvent face of the 40S subunit. Ribosomal proteins are shown in green and 18S rRNA in grey. (PDB 3JAP)

**(B)** Intersubunit face of the 40S subunit. Ribosomal proteins are shown in green and 18S rRNA in grey. (PDB 3JAP)

**(C)** Path of the mRNA through the ribosome. mRNA enters the ribosome above the shoulder and threads through a channel between the 40S head and body, winding around the neck through the A, P, and E decoding sites. When an AUG is present in the P site, the coding region 3' of the start codon occupies the 12 nucleotide entry channel and A site. The +4 nucleotide sits in the A site nearest the P site. The first three upstream context nucleotides (-1 to -3) occupy the E site, and the context further upstream is located in the 12 nt exit channel.

**(D)** Intersubunit face of the 60S subunit, showing the central protuberance (cp), L1 stalk (L1), and L7/L12 stalk (St). (Murray, 2016)

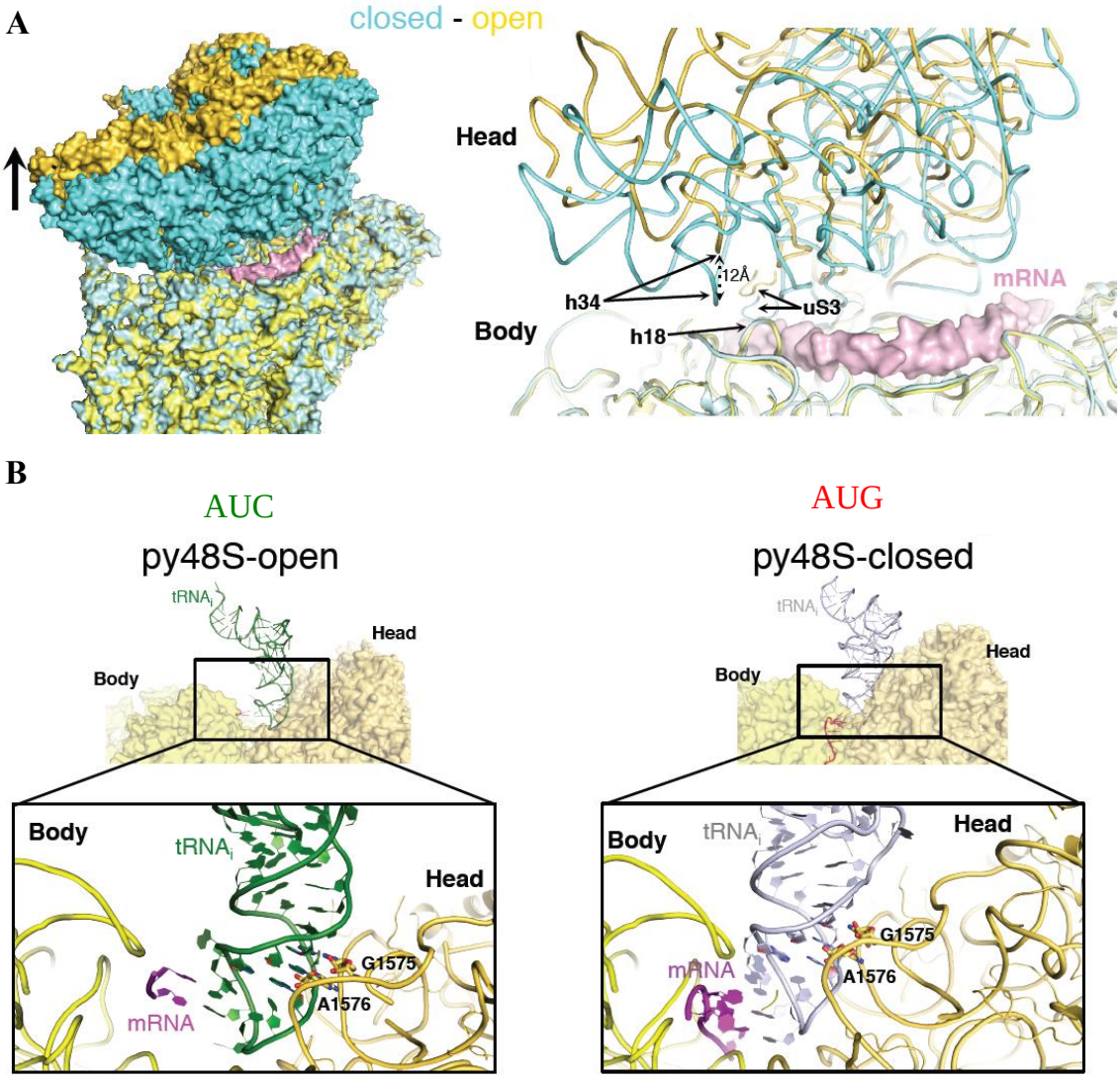

A
closed - open
Head
12Å
h34
uS3
mRNA
Body
h18
B
AUC
py48S-open
tRNAi
Head
Body
Body
tRNAi
Head
G1575
A1576
mRNA
AUG
py48S-closed
tRNAi
Head
Body
Body
tRNAi
Head
G1575
mRNA
A1576

**Figure 1.7.3. Conformational changes in 40S sturcure following AUG recognition.**

**(A)** py48S-open (yellow) and py48S-closed (blue) structures of the ribosome demonstrate a 12 Å closure of the latch between h18 and h34, accompanied by a downward movement of the 40S head and constriction of the mRNA entry channel.

**(B)** Rearrangement of the P site from a relatively open structure (py48S-open, left) to a constricted one (py48S-closed, right) is accompanied by the formation of contacts between $tRNA_i$ and the 40S body that are absent in the open state. (Llacer, 2015)

TC
eIF1
eIF1A
Open PIC
GTP
GDP+ Pi
eIF1A
A
P
E
UUG
eIF1
Open, Scanning PIC
eIF1
GTP
GDP
Pi
eIF1A
A
P
E
AUG
Closed, Arrested PIC

**Figure 1.7.4. AUG recognition is accompanied by transition of the PIC from an open to a closed state.**

(i) Binding of eIF1 and eIF1A to the 40S small ribosomal subunit promotes the open conformation, to which TC rapidly binds. (ii) The open conformation facilitates mRNA loading and scanning, during which the mRNA is inspected for an AUG codon. In this state, Met-tRNA$_i$ is not bound fully in the P site ($P_{OUT}$). Hydrolysis of eIF2-bound GTP may occur during scanning, but $P_i$ release is blocked by the presence of eIF1. (iii) Following AUG recognition, dissociation of eIF1 allows $P_i$ release. The 40S subunit transitions to a closed, arrested state that is not conducive to continued scanning, featuring the constricted mRNA entry channel and P site shown in Fig. 1.7.3. Met-tRNA$_i$ is bound more deeply in the P site ($P_{IN}$), resulting in tighter binding, and the anticodon is base-paired to the AUG start codon.

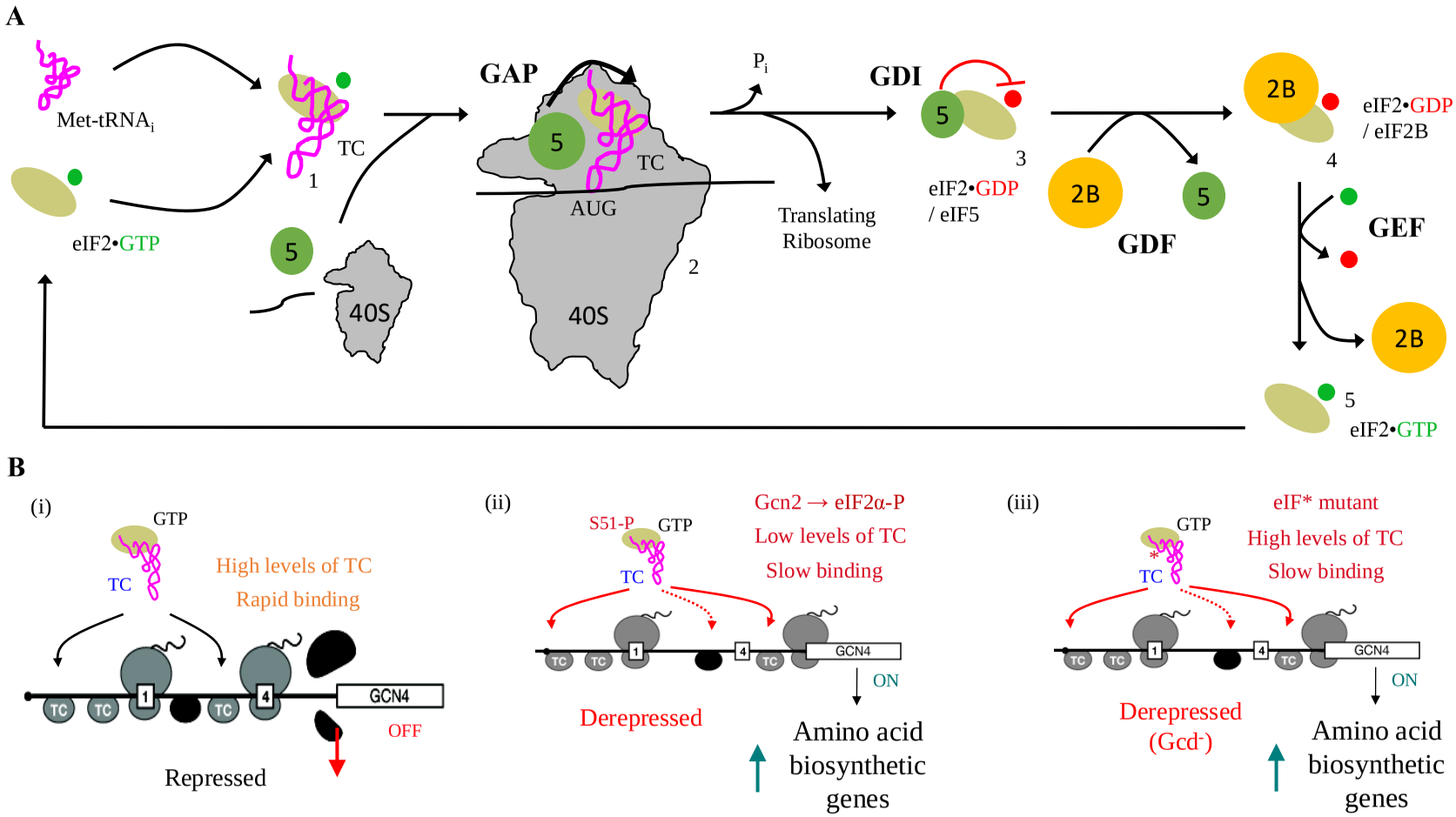

A
Met-tRNA$_i$
eIF2•GTP
TC
1
5
40S
GAP
5
TC
AUG
2
40S
P$_i$
Translating Ribosome
GDI
5
3
eIF2•GDP / eIF5
2B
5
GDF
2B
eIF2•GDP / eIF2B
4
GEF
2B
5
eIF2•GTP
B
(i)
GTP
TC
High levels of TC
Rapid binding
TC
TC
1
TC
4
GCN4
OFF
Repressed
(ii)
S51-P
GTP
TC
Gcn2 → eIF2α-P
Low levels of TC
Slow binding
TC
TC
1
4
TC
GCN4
ON
Derepressed
Amino acid biosynthetic genes
(iii)
GTP
*
TC
eIF* mutant
High levels of TC
Slow binding
TC
TC
1
4
TC
GCN4
ON
Derepressed (Gcd$^-$)
Amino acid biosynthetic genes

**Figure 1.7.5. eIF2 cycle impacts cellular availability of TC.**

**(A)** eIF2•GTP binds Met-tRNA$_i$ to form TC (1), which participates in translation initiation by binding to the PIC. Upon start codon recognition, hydrolysis of the GTP on eIF2 is stimulated by the GTPase activating protein (GAP) activity of eIF5 (2). $P_i$ is released and eIF2•GDP dissociates from the PIC in the company of eIF5, which acts as a GDP dissociation inhibitor (GDI) and delivers eIF2 to eIF2B (3). eIF2B displaces eIF5 on eIF2, acting as a GDI dissociation factor (GDF), resulting in an eIF2•GDP/eIF2B complex (4). eIF2B acts as a guanine exchange factor (GEF), facilitating nucleotide exchange to eIF2•GTP. eIF2, now charged with GTP, can begin a new cycle by forming a new TC (5).

**(B)** *GCN4*, a stress response gene responsible for translational control in response to nutrient stress in yeast, contains four upstream open reading frames (uORFs). The first uORF of *GCN4* (uORF1) is short and contains sequences that promote continued ribosomal scanning. As a result, the scanning PIC must acquire new TC before it can reinitiate. In healthy, WT cells (i), TC loads rapidly onto the PIC, which then reinitiates at one of the other *GCN4* uORFs. After translating this uORF, the ribosome dissociates normally, preventing translation of the *GCN4* coding sequence. When phosphorylation of eIF2 in response to stress decreases TC levels (ii), or when a mutation prevents rapid TC loading (iii), the ribosome scans past uORFs 2, 3, and 4 before TC binds, allowing translation of the coding sequence and turning on *GCN4* expression.

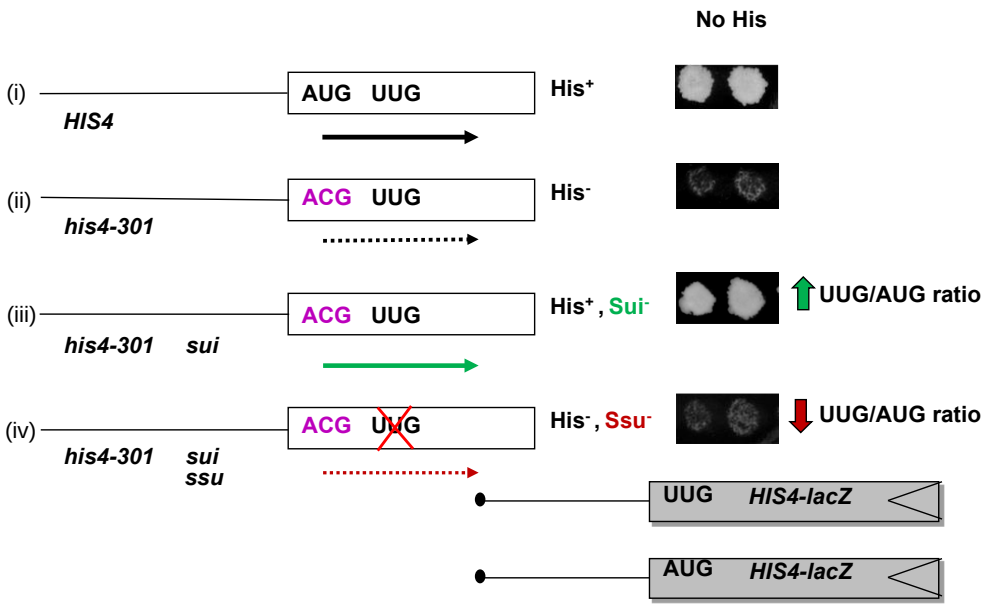

No His
(i)
HIS4
AUG UUG
His+
(ii)
his4-301
ACG UUG
His-
(iii)
his4-301 sui
ACG UUG
His+, Sui-
UUG/AUG ratio
(iv)
his4-301 sui ssu
ACG UUG
His-, Ssu-
UUG/AUG ratio
UUG HIS4-lacZ
AUG HIS4-lacZ

**Figure 1.7.6. Strains containing *his4-301* used to assay initiation fidelity.**

Depiction of the WT *HIS4* gene and *his4-301*, in which the AUG start codon has been mutated to ACG. (i) Strains containing WT *HIS4* can produce histidine and have a His+ phenotype. (ii) Otherwise WT strains containing *his4-301* are not able to produce *HIS4* and are unable to grow on media lacking histidine. (iii) The introduction of a $Sui^-$ mutation in an initiation factor decreases the fidelity of start codon selection, allowing initiation at a downstream UUG, producing a $His^+$ phenotype. Growth is seen on media lacking histidine, and when assaying paired *HIS4-lacZ* reporters (below), the UUG:AUG ratio is increased. (iv) An $Ssu^-$ mutation suppresses the $Sui^-$ phenotype of a mutation in another factor, abolishing the ability to initiate at UUG downstream and restoring the $His^-$ phenotype of *his4-301*. There is a suppression of growth on –His media, and a decrease in the UUG:AUG ratio.

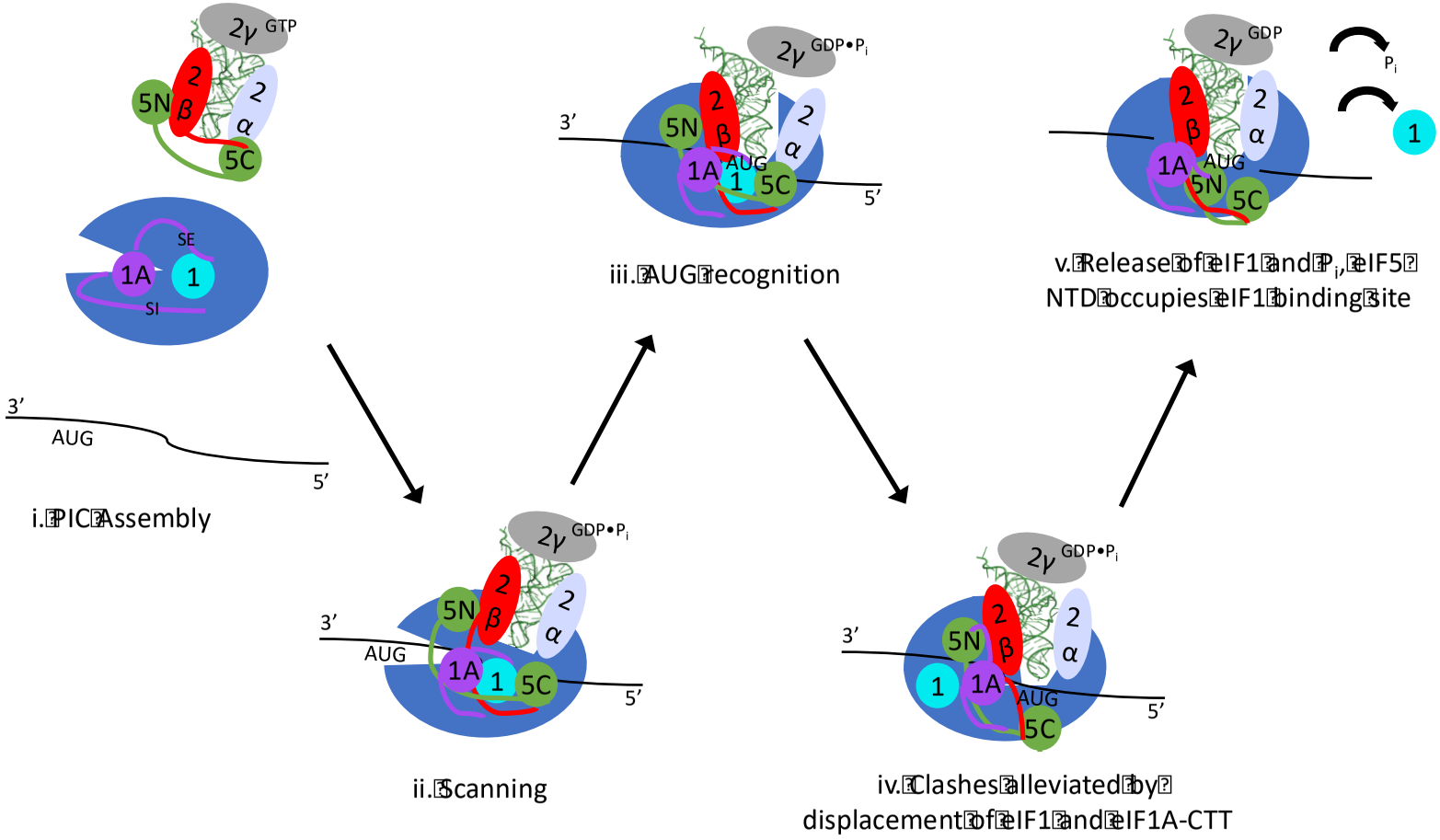

2γ GTP
2
β
2
α
5N
5C
SE
1A
1
SI
3′
AUG
5′
i. PIC Assembly
2γ GDP•Pi
2
β
2
α
5N
1A
AUG
1
5C
3′
5′
iii. AUG recognition
2γ GDP
Pi
2
β
2
α
1A
AUG
5N
5C
1
v. Release of eIF1 and Pi, eIF5 NTD occupies eIF1 binding site
2γ GDP•Pi
2
β
2
α
5N
1A
1
5C
3′
AUG
5′
ii. Scanning
2γ GDP•Pi
2
β
2
α
5N
1
1A
AUG
5C
3′
5′
iv. Clashes alleviated by displacement of eIF1 and eIF1A-CTT

**Figure 1.7.7. Model of start codon recognition.**

(i.) eIF1 and scanning enhancer elements in the CTT of eIF1A promote an open conformation of the 40S SSU to which TC binds, along with initiation factors of the MFC, including eIF5. The resultant 43S PIC loads onto the 5' end of mRNA. (ii.) The open conformation of the PIC facilitates scanning in the 5'-to-3' direction, during which the mRNA is sampled for complementarity to the anticodon of Met-tRNA$_i$. The GTP on eIF2 may be hydrolyzed to GDP•P$_i$ during scanning, but the release of inorganic phosphate is blocked by the presence of eIF1. (iii.) When an AUG is recognized, a conformational change of the 40S complex from an open to a closed state results in a constricted mRNA entry channel and P site. This change is accompanied by deformation of the tRNA as it binds deeper in the P site, facilitating base pairing of the anticodon to AUG. In its new position, Met-tRNA$_i$ clashes with Loops 1 and 2 of eIF1 and the CTT of eIF1A. (iv.) Clashes with Met-tRNA$_i$ in the closed state are alleviated by deformation of eIF1 Loops 1 and 2, and ultimately, displacement of eIF1 to the 40S platform. The eIF1A CTT dissociates from the eIF5 CTD and moves from the P site to interact with the eIF5 NTD. This results in a strong interaction between the eIF5 CTD and the NTT of eIF2β. (v.) Dissociation of eIF1 from the PIC and the interaction of the eIF1A CTT with the eIF5 CTD allow the release of P$_i$ from eIF2. The NTD of eIF5 is repositioned to the eIF1 binding site, preventing reassociation of eIF1.

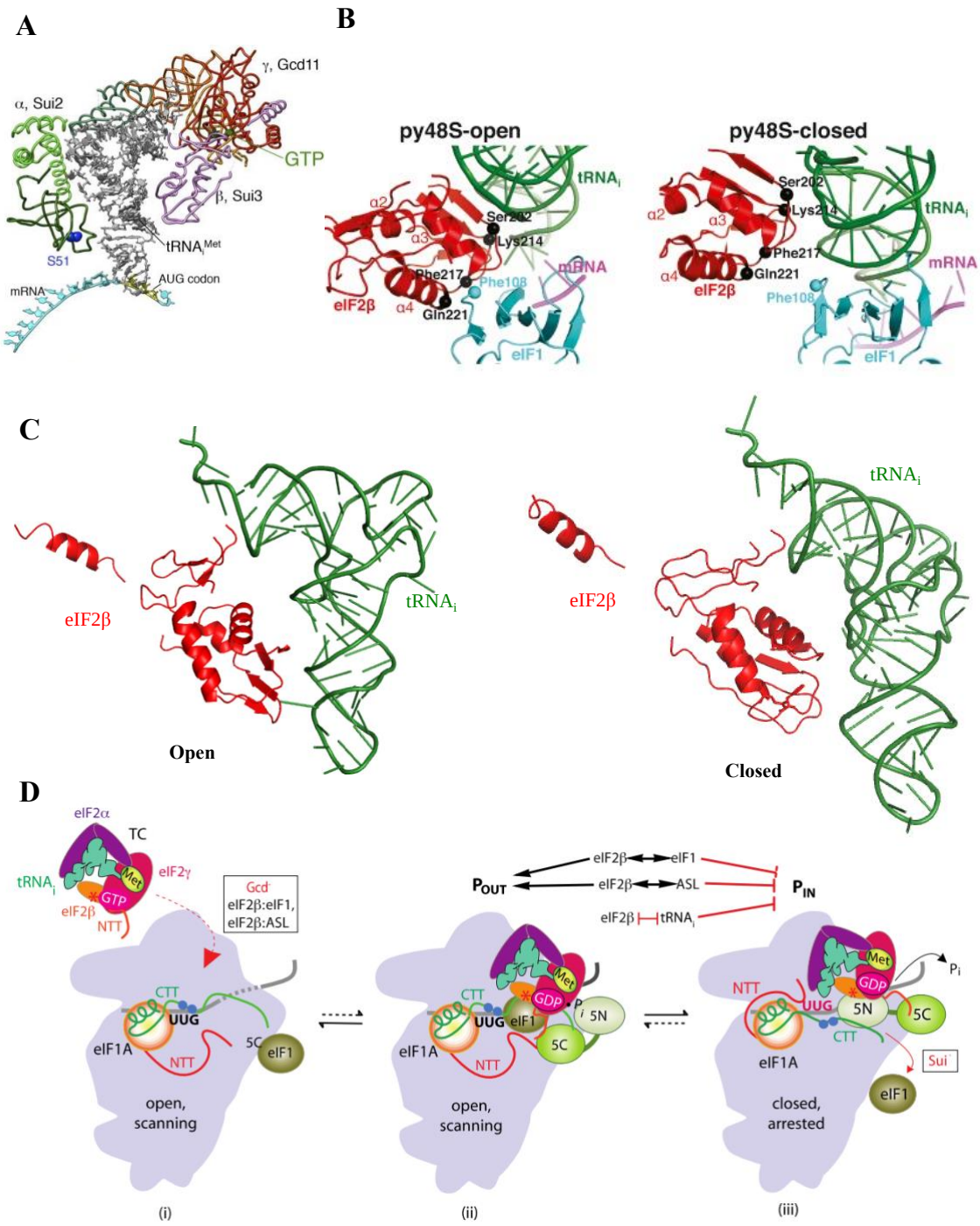

A
γ, Gcd11
α, Sui2
GTP
β, Sui3
tRNA$_i^{Met}$
S51
AUG codon
mRNA
B
py48S-open
tRNA$_i$
α2
Ser262
α3
Lys214
Phe217
mRNA
Phe108
eIF2β
α4
Gln221
eIF1
py48S-closed
Ser262
Lys214
tRNA$_i$
α2
α3
Phe217
α4
Gln221
mRNA
eIF2β
Phe108
eIF1
C
eIF2β
tRNA$_i$
Open
eIF2β
tRNA$_i$
Closed
D
eIF2α
TC
tRNA$_i$
Met
eIF2γ
GTP
eIF2β
NTT
Gcd$^-$
eIF2β:eIF1,
eIF2β:ASL
CTT
UUG
eIF1A
5C
eIF1
NTT
open,
scanning
(i)
P$_{OUT}$
eIF2β
eIF1
eIF2β
ASL
eIF2β
tRNA$_i$
P$_{IN}$
Met
GDP
P$_i$
CTT
UUG
eIF1
5N
5C
eIF1A
NTT
open,
scanning
(ii)
Met
GDP
P$_i$
NTT
UUG
5N
5C
CTT
eIF1A
Sui$^-$
eIF1
closed,
arrested
(iii)

**Figure 1.7.8. eIF2 stabilizes the open PIC while preventing transition to the closed state.**

**(A)** Structure of eIF2 bound in ternary complex with Met-tRNA$_i$ (grey) and bound to mRNA (teal). eIF2γ (orange) acts as a backbone for the complex, binding eIF2α (green), with the phosphorylation site S51 marked in blue, and eIF2β (purple). (pdb 3JAP, Llacer, 2015; presented in Dever, 2016)

**(B)** Structure of eIF2β (red) binding Met-tRNA$_i$ (green) and eIF1 (cyan) in the py48S-open and py48S-closed complex. The contacts between eIF2β and eIF1 are lost in the closed complex. (pdb 3JAP, Llacer, 2015)

**(C)** eIF2β (red) forms differential contacts with Met-tRNA$_i$ (green) in the py48S-open and py48S-closed complexes. Contacts between eIF2β and the anticodon stem loop are lost in the closed complex. (pdb 3JAP, Llacer, 2015)

**(D)** Model depicting the dual role of eIF2β in start codon selection. (i) eIF1 promotes the open complex of the PIC to which WT TC rapidly loads. Mutations that disrupt the eIF2β:eIF1 interface or eIF2β contacts with the ASL cause TC loading defects, confering Gcd$^-$ phenotypes. (ii) Once bound, eIF2β promotes the open conformation by helping to anchor eIF1 onto the complex and by stabilizing binding of Met-tRNA$_i$ through its contacts with the ASL. (iii) At the same time, eIF2β inhibits transition to the closed state in the absence of an appropriate start codon. Preventing eIF1 dissociation and stabilizing Met-tRNA$_i$ both help to prevent this transition. Further, eIF2β forms steric and electrostatic clashes with the closed state position of Met-tRNA$_i$, preventing its rearrangement to the closed state until an AUG is recognized.

**A**

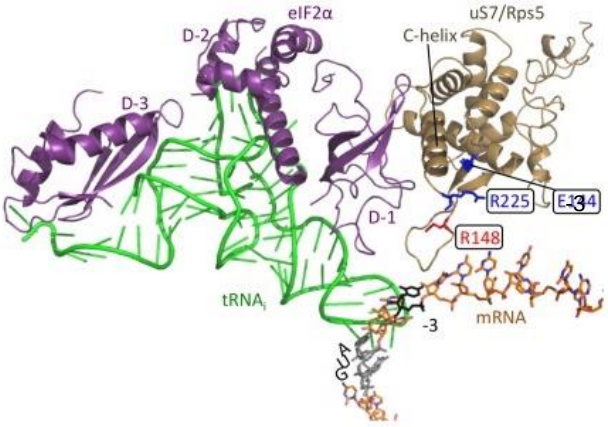
**Consensus sequence:**

**5' Cap**

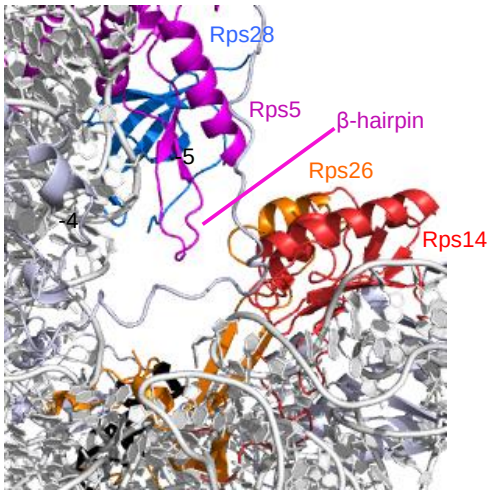
**Mammals (Kozak)** A/G$^{-3}$C$^{-2}$C$^{-1}$**AUG**G$^{+4}$

≥32 nts ***S. cerevisiae*** A$^{-3}$A$^{-2}$A/G$^{-1}$**AUG**A$^{+4}$ **mRNA**

**B**




**C**




E site and Exit Channel,
Front View

**D**

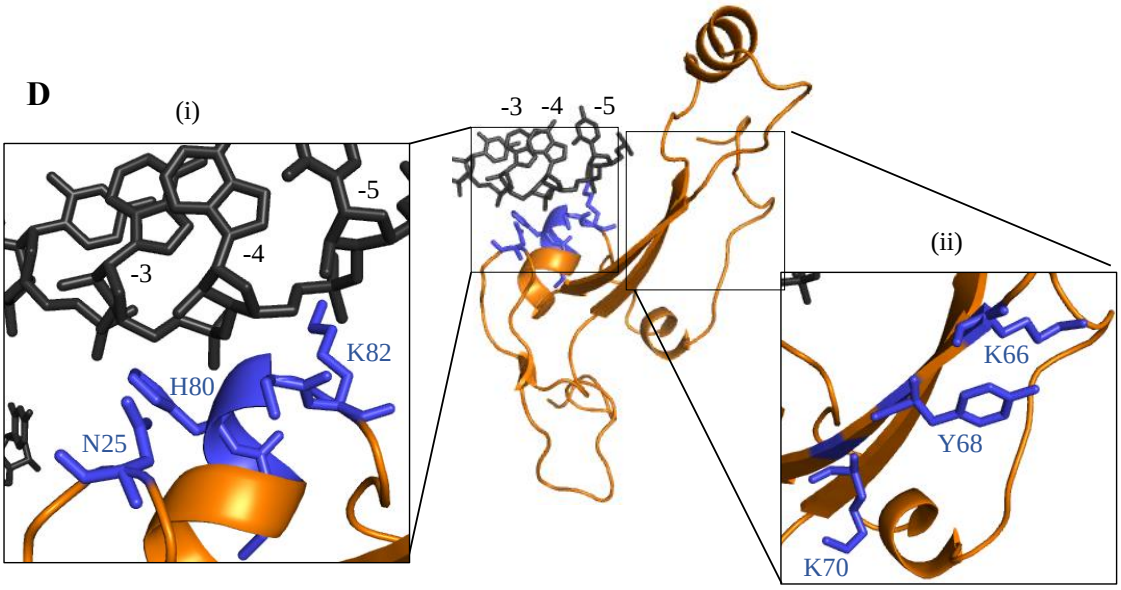


**Figure 1.7.9. Rps26/eS26 contacts critical nucleotides of the Kozak sequence.**

**(A)** Kozak consensus sequence in yeast and mammals. The presence of purines at the -3 and +4 positions, marked in red, have been determined to have the greatest impacts on start codon selection.

**(B)** Rps5/uS7 (gold) and eIF2α (purple) contact mRNA (orange) at the important -3 context nucleotide (black) in a structure of the py48S PIC. In addition, the C-helix of Rps5 interacts with eIF2α -D1. (pdb 3J81, Hussain, 2014; presented in Visweswaraiah, 2017)

**(C)** Four ribosomal proteins, Rps5/uS7 (pink), Rps14 (red), Rps26/eS26 (orange), and Rps28/eS28 (blue) occupy the exit channel, with the potential to contact mRNA context nucleotides upstream of a start codon.

**(D)** Depiction of py48S PIC showing contacts between Rps26/eS26 (orange) and mRNA (black). Proposed contacts based on structural evidence (i) and hydroxyl radical crosslinking (ii) are colored in blue. (pdb 3JAP, Llacer, 2015)

# 2. Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex


Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex
Jose L. Llácer,[1,4] Tanweer Hussain,[1,4] Laura Marler,[2] Colin Echeverría Aitken,[3] Anil Thakur,[2] Jon R. Lorsch,[3,5] Alan G. Hinnebusch,[2,5] and V. Ramakrishnan[1,*]

Mol Cell. 2015 Aug 6; 59(3): 399–412.

## 2.1 ABSTRACT

Translation initiation in eukaryotes begins with the formation of a pre-initiation complex (PIC) containing the 40S ribosomal subunit, eIF1, eIF1A, eIF3, ternary complex (eIF2-GTP-Met-tRNA$_i$), and eIF5. The PIC, in an open conformation, attaches to the 5' end of the mRNA and scans to locate the start codon, whereupon it closes to arrest scanning. We present single particle cryo-electron microscopy (cryo-EM) reconstructions of 48S PICs from yeast in these open and closed states, at 6.0 Å and 4.9 Å, respectively. These reconstructions show eIF2β as well as a configuration of eIF3 that appears to encircle the 40S, occupying part of the subunit interface. Comparison of the complexes reveals a large conformational change in the 40S head from an open mRNA latch conformation to a closed one that constricts the mRNA entry channel and narrows the P site to enclose tRNA$_i$, thus elucidating key events in start codon recognition.

## 2.2 INTRODUCTION

In the first stage of eukaryotic translation initiation, the 40S ribosomal subunit and translation initiation factors eIF1, eIF1A, and eIF3 form a complex that facilitates loading of methionyl initiator tRNA ($tRNA_i$) onto the 40S subunit as a ternary complex (TC) with eIF2-GTP. eIF5, a GTPase activating protein (GAP) for eIF2, is thought to bind to this complex along with TC. The 43S preinitiation complex (PIC) thus formed binds to the capped 5′ end of mRNA in collaboration with the eIF4F complex, which consists of the cap-binding protein eIF4E, scaffolding protein eIF4G, and RNA helicase eIF4A. This 48S PIC, in an open conformation with $tRNA_i$ not fully engaged in the P site ($P_{OUT}$), then scans along mRNA. During the scanning process, GTP bound to eIF2 is hydrolyzed but the dissociated phosphate ($P_i$) is not released. Recognition of the start codon in the P site precipitates transition to a scanning-arrested, closed PIC with $tRNA_i$ accommodated in the P site ($P_{IN}$). This rearrangement triggers release of eIF1 and resultant dissociation of $P_i$ (12, 319-321).

One consequence of this process is that the mRNA cannot be threaded into the 40S subunit, because eIF4F is bound to the 5′ end, and hence must be loaded laterally into the mRNA channel. In the empty 40S subunit, this channel is closed because interactions between helix h34 in the head and h18 in the body form a latch (127) that must open to allow initial loading of mRNA on the 40S subunit. It is thought that eIF1 and eIF1A promote the formation of an open conformation of the 40S subunit conducive to scanning (151, 322). Moreover, a low-resolution cryo-EM reconstruction (22 Å) of a 40S•eIF1•eIF1A complex suggested that eIF1 and eIF1A unlock the latch, but the factors themselves could not be seen in this structure (127). In contrast, all recent structures, including those with eIF1 and eIF1A, show the latch in a closed conformation (91, 92, 109, 110, 138, 195). A complete structure of the open conformation, in

which the positions of the various factors are visible, would shed light on the mechanism of initial mRNA loading, scanning, and start codon recognition, as well as the roles played by each factor in these events.

Apart from capturing only the closed-latch conformation, previous structures are also incomplete in other ways. A recent 48S complex (py48S) from our groups shows details of the interaction of eIF1 and eIF1A with the ribosome and $tRNA_i$ during start-codon recognition, but it lacks density for most of the β subunit of eIF2 and all of eIF3 (110). eIF3 is the largest and most complex of the eIFs, and is involved in nearly every aspect of initiation (250, 323). Mammalian eIF3 consists of 13 subunits and has a molecular weight of ~800 kDa, whereas eIF3 in *Saccharomyces cerevisiae* and closely related yeasts is a ~395 kDa complex of five essential subunits (eIF3a, eIF3b, eIF3c, eIF3g, and eIF3i) and a sixth non-essential and substoichiometric subunit, eIF3j (145, 250). Crystal structures of parts of the various subunits of eIF3 from yeast have recently been solved (12, 324), and a cryo-EM structure of 40S•eIF1•eIF1A•eIF3 is now available (114). Domains of eIF3 were also modeled in a moderate resolution cryo-EM reconstruction of a partial mammalian 43S complex (hereafter pm43S) that lacked eIF1, eIF1A, eIF5, and mRNA (109). All known structures of eIF3 bound to the 40S, as well as crosslinking studies, indicate binding on the solvent surface of the subunit (109, 114, 324, 325). And yet eIF3 has been shown to interact with eIF1, TC, and eIF5, all of which bind to the intersubunit surface of the 40S (12). How eIF3 interacts with components bound near the P site while itself binding to the opposite face of the 40S subunit remains a key question in understanding the mechanism of translation initiation.

Here we have addressed these questions by determining a cryo-EM reconstruction at 6.0 Å resolution of a partial yeast 48S complex (with mRNA containing a near-cognate AUC in

lieu of an AUG codon) that reveals an open, scanning-competent state with $tRNA_i$ not fully engaged in the P site, hereafter referred to as "py48S-open." A second complex elucidated at 4.9 Å resolution contains mRNA with an AUG codon that presents a closed conformation with the latch closed, entry channel constricted, and $tRNA_i$ locked into the $P_{IN}$ state, hereafter referred to as "py48S-closed." These structures show clear density for eIF1, eIF1A, mRNA, the entire TC, including previously unseen eIF2β, as well as segments of eIF3. The py48S-open complex shows an open latch conformation, expanded entry channel and widened P site, suggesting mechanisms for loading and scanning of mRNA and is markedly different from a pm48S reconstructed at 11.6 Å (109). Comparison with py48S-closed illuminates the structural changes that occur within the 40S subunit, TC, and other eIFs during the transition from the open to closed state of the PIC that should arrest scanning and lock $tRNA_i$ into the P site and highlights the importance of the 40S head conformation and roles of eIFs in stabilizing the two states. We also observe portions of eIF3 on the subunit-joining interface of the 40S subunit in both complexes, showing how eIF3 can contact TC and eIF1 near the P site while remaining bound to the solvent face of the 40S subunit.

## 2.3 MATERIALS AND METHODS

### 2.3.1 Electron Microscopy

Data were collected on an FEI Titan Krios microscope operated at 300 kV under low-dose conditions (27 $e^{-}/Å^2$) using a defocus range of 1.8–3.2 μm. Images were recorded on a Falcon II detector at a calibrated magnification of 104,478 (yielding a pixel size of 1.34 Å). An

in-house system was used to intercept the videos from the detector at a speed of 16 frames/s exposures. Micrographs that showed noticeable signs of astigmatism or drift were discarded.

### 2.3.2 Analysis and Structure Determination

Particles were picked using RELION (326). Contrast transfer function parameters for the micrographs were estimated using CTFFIND3 (327). 2D class averaging, 3D classification, and refinements were done using RELION (326).

Statistical movie processing was done (328) to improve resolution of all reconstructions. Resolutions reported are based on the gold-standard FSC = 0.143 criterion (329). Local resolution was estimated using RESMAP (330). All maps were further processed for the modulation transfer function of the detector and sharpened by applying negative B factors ($-20$ Å$^2$ for py48S-open and $-119$ Å$^2$ py48S-closed; estimated as in (331).

### 2.3.3 Model Building and Refinement

The atomic model of py48S (PDB: 3J81) was placed into density by rigid-body fitting using Chimera (332). Further model building was done in Coot (333). For py48S-open, the body and head of the 40S were independently placed. For eIF2β, models from its archaeal counterpart were employed (PDB: 3CW2, 2D74). Wild-type $tRNA_i$ was used from PDB: 1YFG for initial rigid-body fitting into its corresponding density in the py48S-open complex. Model building and

refinement were carried out using Coot and Refmac (334) (see Supplemental Experimental Procedures, Section 6.3). All figures were generated using PyMOL (DeLano, 2006) or Chimera.

## 2.4 RESULTS

### 2.4.1 Formation and Overview of Structures

Yeast py48S-closed was assembled as described previously (110) using an unstructured, uncapped mRNA with an AUG codon; 40S subunits from yeast *Kluyveromyces lactis*; and factors eIF1, eIF1A, eIF3, eIF5, and TC from *S. cerevisiae*. We used the U31:A39 variant of $tRNA_i$ that stabilizes the $P_{IN}$ state (203) to promote formation of the 48S in the $P_{IN}$ state, but wild-type eIF2 rather than the Sui3-2 variant of eIF2 (282) used previously (110). We similarly combined 40S with eIF1, eIF1A, eIF3, TC, and mRNA to generate the py48S-open complex. However, to shift the equilibrium from the $P_{IN}$ state, wild-type $tRNA_i$ was used, the AUG codon was replaced with near-cognate AUC, eIF5 was omitted as it shifts the equilibrium toward $P_{IN}$ state (211, 222), and recombinant eIF3 expressed in *E. coli* (and thus free of eIF5) was used instead of native eIF3 (see Supplemental Experimental Procedures, Section 6.3).

The structures for py48S-closed and py48S-open were determined to overall resolutions of 4.9 Å and 6.0 Å, respectively (Figures 2.10.1, 6.1.1, and 6.1.2; Table 6.1.1; Movies S1 and S2). Like py48S (110) the local resolution and density is best for the core of 40S and ligands directly attached to it (Figure 6.1.3 and Table 6.1.1). Large data sets and extensive 3D classification were required to obtain PIC classes with eIF1, eIF1A, eIF3, TC (including eIF2β), and mRNA all bound (Figure 6.1.1; see Supplemental Experimental Procedures, Section 6.3).

These classes comprised 1.8% of the total for py48S-closed and 1.0% for py48S-open. The majority of particles were deficient in one or more factors as a result of the characteristic dissociation of factors on the EM grid (Figure 6.1.1; See Supplemental Experimental Procedures, Section 6.3).

We observe density in both complexes for subunits of eIF3, as well as for eIF1, eIF1A, mRNA, and all three subunits of TC. Interestingly, density for components of eIF3 appears in the subunit-joining interface, where it interacts with other eIFs. In the solvent interface, the density for eIF3 density is weaker, especially for py48S-open (discussed later). We observe density for eIF2β in both complexes (Figures 2.10.1 and 6.1.2). The py48S-closed (4.9 Å) is globally similar to our previously reported py48S (4.0 Å) (110) (RMSD of 0.86 Å for 33,178 atoms of 18S) except for the additional densities for eIF2β and eIF3. The use of large data sets may have allowed us to isolate a class that includes eIF2β and eIF3, which would have been missed earlier in py48S. However, it is also possible that the use of WT eIF2 instead of Sui3-2 variant may have resulted in observation of these eIFs.

The large data set and extensive classification also enabled us to determine structures of a 40S•eIF1•eIF1A complex at 3.5 Å resolution (Figures 6.1.1 and 6.1.4 A) and a 40S•eIF1•eIF1A•TC complex, representing a partial 43S PIC (py43S), at 15 Å resolution (Figures 6.1.1 and 6.1.4 B), which has a conformation similar to pm43S (109).

### 2.4.2 Altered Conformation of the 40S Head in the Open Conformation of py48S

Whereas the orientations of the 40S body are similar in both py48S-open and py48S-closed, the two structures differ in the conformation of the 40S head (Figure 2.10.2; Movie S3).

While the orientation of the head in py48S-closed is similar to that in py48S reported earlier (110), in py48S-open, there is a remarkable upward movement of the head away from the body (Figures 2.10.2 A - B), in a different direction from that distinguishing the 40S•eIF1•eIF1A complex from empty 40S or py48S (Movie S4). This head movement from py48S-closed to py48S-open is accompanied by a 7–8 Å change in the pitch of h28 (Figures 2.10.2 C - D) and a repositioning of the β-hairpin of uS5 that contacts h28 (Figure 2.10.2 C). This helix constitutes the "neck" of the 40S that connects the head to the body, and is compressed in py48S-closed but relaxed in py48S-open. Interestingly, mutations in this region of h28 (A1151, A1152, and U1627; *S. cerevisiae* numbering throughout the manuscript) lead to a $Gcd^-$ phenotype, indicating less stable TC binding to the PIC (197).

This conformation of the head in py48S-open throws open the mRNA channel latch and widens the channel, particularly at the entry channel side near the A site (Figure 2.10.2 B). Helix 34 and associated elements move away from h18 to open the latch. The py48S-open structure reveals both the upward shift of the 40S head and open-latch conformation, thus providing insights into changes involved during key steps of initiation.

The 40S head is also moved upward in py48S-open compared to pm43S (109) (Figure 6.1.4 C). In contrast, the head conformations of the pm43S and py48S-closed are more similar (Figure 6.1.4 D). Note that pm43S lacks mRNA, and densities for eIF1, eIF1A, and eIF2β were missing in the reconstruction. In py43S (which most closely mimics pm43S), the latch is again closed and the head orientation is almost identical to that of pm43S (Figure 6.1.4 E). The positions of TC in pm43S and py43S are also similar (Figure 6.1.4 E); however, eIF1 and eIF1A densities are clearly seen in our py43S map (Figure 6.1.4 B). Thus, the presence of TC with eIF3 in pm43S, or TC with eIF1 and eIF1A in py43S here, both lead to a similar

orientation of the head with a closed latch and may thus represent a state prior to the binding of mRNA.

### 2.4.3 A Widened P Site and Altered Orientation of $tRNA_i$ in py48S-Open

Interestingly, the P site in py48S-open is widened compared to that of py48S-closed (Figures 2.10.3 A - B), lacking interactions between $tRNA_i$ and the 40S body that occur in the closed complex (Figures 2.10.3 A - B). As a result of the altered head position in py48S-open, the $tRNA_i$ adopts a previously unobserved modified P/I orientation, which we call sP/I for "scanning P/I" conformation (Figure 6.1.5 A; Movie S3). Nevertheless, the $tRNA_i$ maintains the same contacts with the head in both complexes, which ensures that the conserved GC base pairs in the anticodon stem-loop (ASL) of $tRNA_i$ are recognized by rRNA residues G1575 and A1576 in both py48S-closed and py48S-open (Figures 2.10.3 A - B).

While the tip of the ASL of $tRNA_i$ is deep within the P site in both the open and closed complexes, it is displaced laterally from the body by ~7 Å in the py48S-open (Figure 2.10.3 C), owing to both the widened P site and altered head orientation. Superimposing the head in the open and closed structures shows that the positions of the ASL relative to the head are very similar in both (Figures 2.10.3 D - E); the ASL essentially moves with the head during the open-to-closed transition. In contrast, in pm43S (109), the tip of the ASL is not deep in the P site (110) and thus does not track with the head movement.

Although the ASL tracks with the head as it moves from the open to closed conformation, the positions of the $tRNA_i$ acceptor arm in both structures are superimposable relative to the

body (Figure 2.10.3 C). As the ASL remains in contact with the head, it must bend in going from the open to closed state, which allows the $tRNA_i$ to maintain codon-anticodon interaction during the transition that would otherwise clash with the 40S body (Figure 6.1.5 B and Movie S3). The bent ASL in py48S-close is similar to that observed earlier in py48S (110) and pm48S (195), and allows base pairing with the codon in $P_{IN}$ state.

### 2.4.4 Path of mRNA in the Two Structures

In our previous py48S complex (110), mRNA was observed throughout the mRNA channel. In contrast, the py48S-closed here shows density for mRNA mainly in the exit channel (Figure 6.1.2 A). Strikingly, in py48S-open, the mRNA entry channel is widened, which, along with the open latch, produces a conformation that should allow single-stranded mRNA to be slotted directly into the mRNA-binding channel (Figures 2.10.2 A - B). We observe discontinuous densities for mRNA, mainly in the P and E sites of py48S-open, including density consistent with base pairing between the A and U of the AUC codon and the U and A of the anticodon (Figure 6.1.2 B, right). Incidentally, the mRNA also has another AUC codon, but because it is only 3 nt away from 5′ end, it appears not to be involved in recognition. The P-site codon has moved in concert with the ASL and 40S head such that base pairing is maintained in py48S-open, despite the $tRNA_i$ not yet having been fully accommodated in the eP/I′ configuration (Figures 2.10.3 E and 6.1.5 B). The minimal density for mRNA in py48S-open suggests that, probably as a result of the widened mRNA channel, the mRNA has minimal contact with the ribosome apart from base pairing with the anticodon, which should facilitate scanning. Upon AUG recognition, head repositioning stabilized by interaction of the N-terminal

tail (NTT) of eIF1A with the codon:anticodon duplex (discussed below), and latch closure, will narrow the entry channel to fix the mRNA and arrest scanning.

### 2.4.5 Changes in eIF1and eIF1A between the Closed and Open States of py48S

We observe eIF1and eIF1A clearly in both the open and closed PICs (Figures 2.10.1, 6.1.2 A, and 6.1.2 B), but with marked changes in their conformations between the two states. The overall conformation of eIF1A is similar in the two complexes, but the NTT of eIF1A interacts with the anticodon-codon duplex only in py48S-closed, as in our previous py48S without eIF3 (110)—consistent with its role in promoting recognition of a cognate start codon (Figure 2.10.4 A). In contrast, the NTT is disordered in py48S-open (Figure 2.10.4 A), and in the 40S•eIF1•eIF1A complex (Figure 6.1.4 A), as in all other reported PICs (110, 138, 195). There is no distinct density for the C-terminal tail (CTT) of eIF1A in any of the closed complexes as expected from previous hydroxyl radical cleavage studies that show that eIF1A-CTT interferes with the P-site tRNA (211, 220). The CTT of eIF1A is also not modeled in py48S-open because of lack of clear unambiguous density.

The overall position of eIF1 in the two complexes is also similar (Figure 2.10.4 B), but β-hairpins 1 and 2 of eIF1 are positioned differently. Their orientations in py48S-open resemble those observed in the 40S•eIF1•eIF1A complex, with no steric clash with $tRNA_i$. In py48S-closed, however, the two β-hairpins are displaced to avoid a clash with the now-accommodated $tRNA_i$ (Figure 2.10.4 B), as seen earlier in py48S (110). An interesting feature in py48S-open is the interaction of β-hairpin 1 with the AUC codon in the P site (Figure 2.10.3 E). As we observe a similar interaction with the AUG codon in py48S-closed (Figure 6.1.5 B), the conformation of

β-hairpin 1 changes between the two states to follow the $tRNA_i$ ASL and P-site codon as they are adjusted during the open-to-closed transition, preserving this interaction (Figure 2.10.4 B).

### 2.4.6 eIF2β Links TC to the 40S Head and Body in the py48S-Open Complex

We observe density for all three subunits of eIF2 in the PIC, including eIF2β (Figures 6.1.2 A and 6.1.2 B). eIF2 is bound primarily to the 40S head in py48S-closed (Figure 2.10.1 A) and, as in py48S (110), eIF2α-D1 (domain 1) and the $tRNA_i$ ASL together attach TC to the head (Movie S3). During the transition from py48S-open to py48S-closed, eIF2α-D1 rotates slightly, thus avoiding a clash with the 40S body (Movie S3). This positions the loop containing Arg55 and Arg57 to enable their interactions with mRNA nucleotides −2 to −3 in the E site (Figures 6.1.6 A and 6.1.6 B), as observed in the py48S. Another consequence of this rotation is that eIF2α-D2 and the D-and T-arms of the $tRNA_i$ are positioned closer to the head in py48S-closed compared to py48S-open (Figures 2.10.3 A - B). Moreover, eIF2α-D2 moves in relation to eIF2α-D1 and interacts closely with the D- and T-loops of $tRNA_i$ (Figure 2.10.5 A; Movie S3). The third domain eIF2α-D3 moves with respect to the acceptor arm of the $tRNA_i$ (Figure 2.10.5 A; Movie S5).

We could model most of eIF2β in both complexes (Figures 2.10.1 and 6.1.2), except for the disordered N-terminal residues 1–125 and the last 20 C-terminal residues of the protein. As in previous archeal βγ complexes (292, 335), eIF2β is tightly attached to eIF2γ by its N-terminal helix α1 (Figures 6.1.2 A and 6.1.2B). Notably, in py48S-open, the helix-turn-helix (HTH) domain of eIF2β binds to eIF1 and eIF1A on the 40S body and to $tRNA_i$ bound to the 40S head, bridging the 40S head and 40S body, without direct interactions with the 40S itself

(Figure 2.10.5 B, upper). These interactions likely stabilize the open conformation during scanning in the absence of a complete codon-anticodon duplex.

During rearrangement to the closed complex, the eIF2β HTH domain is positioned away from eIF1 and eIF1A (Figure 2.10.5 B, lower) and binds to elements of the 40S head. Because of its altered position, the HTH domain also makes contacts with the $tRNA_i$ that are distinct from those occurring in py48S-open (Figure 2.10.5 B, upper). The position of eIF2β in py48S-open would result in a clash with both eIF1 and eIF1A in py48S-closed due to inward movement of the 40S head and body (Figure 6.1.6 C).

The zinc-binding domain (ZBD) of eIF2β is positioned close to the GTP binding pocket of eIF2γ in both complexes (Figure 2.10.5 C), similar to its position in some archaeal βγ complexes (335), although the ZBD itself was disordered in that structure. Because eIF1 is present in both complexes and eIF1 dissociation from the PIC is a prerequisite for $P_i$ release from eIF2•GDP•$P_i$ on AUG recognition (207), it remains unclear whether changes in the interaction of the eIF2β-ZBD with the eIF2γ GTPase center on eIF1 release are involved in $P_i$ release.

In our previously reported py48S structure (110), density for eIF2β was largely absent. This may be because we used the Sui3-2 variant of eIF2 harboring the S264Y substitution in eIF2β to stabilize the closed PIC conformation. Interestingly, S264Y maps to the interface between the eIF2β-ZBD and the eIF2γ GTPase center (Figure 2.10.5 C). As such, it might destabilize this interface and increase mobility of the eIF2β-ZBD, which could disrupt the interactions of eIF2β with eIF1 and eIF1A that are unique to py48S-open (Figure 2.10.5 B). This might explain how the Sui3-2 mutant stabilizes TC binding in the closed conformation (214). Sui3-2 may specifically destabilize the open conformation and shift the equilibrium toward the

closed conformation, thus accounting for its increased utilization of near-cognate start codons *in vivo* (282).

Interestingly, a superposition of eIF2βγ using eIF2γ as a reference shows that the relative orientation of these two subunits is the same in both py48S-open and py48S-closed (Figure 6.1.6 D). Hence, eIF2βγ, along with eIF2α-D3, alters its position relative to $tRNA_i$ and domains 1 and 2 of eIF2α in a concerted manner between the two complexes (Figures 2.10.5 D and 6.1.6 E; Movie S5).

### 2.4.7 Initiation Accuracy *In Vivo* Is Reduced by Disrupting eIF2β Contacts with $tRNA_i$ or eIF1 that Occur Only in py48S-Open

The fact that eIF2β makes interactions with the $tRNA_i$ ASL and eIF1 in py48S-open that are missing or altered in py48S-closed (Figure 2.10.6 A) suggests that these contacts specifically stabilize the open, scanning conformation of the PIC. If so, then substituting residues at these contacts should increase the frequency of initiation at UUG codons by facilitating rearrangement from the open to closed conformations in the absence of a perfect start codon:anticodon duplex in the P site. Supporting this prediction, substitutions of eIF2β Phe217/Gln221 and eIF1 Phe108, residues juxtaposed at the eIF2β/eIF1 interface in py48S-open (Figure 2.10.6 A), substantially increase the UUG:AUG expression ratio for matched *HIS4-lacZ* fusions differing only in the start codon (Figure 2.10.6 B). The eIF1-F108A/F108D substitutions also increase eIF1 abundance (Figure 2.10.6 D), an established indicator of relaxed discrimination against the suboptimal context of the AUG start codon of the (*SUI1*) mRNA encoding eIF1 (212). Thus, eIF1-F108A/F108D facilitate initiation for both a near-cognate start codon and an AUG in poor

context. eIF2β substitutions S202A/K214A also increase UUG initiation (Figure 2.10.6 B) and, consistent with impaired eIF2β interaction with $tRNA_i$, derepress expression of a *GCN4-lacZ* reporter (Figure 2.10.6 C), an *in vivo* indicator of reduced TC assembly or binding to the scanning PIC (295). None of the eIF2β substitutions significantly affect eIF2β abundance (Figure 2.10.6 E) or its assembly with eIF2α/eIF2γ in the eIF2 complex (Figure 2.10.6 F).

### 2.4.8 Placement of eIF3 subunits on both faces of the PIC

We observe density for eIF3 in both py48S-open and py48S-closed (Figures 2.10.1 A - B). At this resolution, we can identify helices, place domains of known structure into the density, and make tentative assignments of previously unobserved segments of eIF3, based on secondary structure predictions (Figures 6.1.2 C, 6.1.2 D, and 6.1.2 E; Supplemental Experimental Procedures, Section 6.3). Because the densities attributable to eIF3 on the subunit interface are similar in both complexes, we describe the appearance of eIF3 only in the higher resolution py48S-closed (Figure 2.10.7).

The two PCI domains of the eIF3a/eIF3c heterodimeric core bind near the left shoulder of the 40S solvent face (Figure 2.10.7 A; Movie S1), as in the yeast 40S•eIF1•eIF1A•eIF3 structure (114). However, in py48S-closed, the PCI domains are displaced laterally, which may reflect a conformational change in eIF3 during different steps of initiation (Figure 2.10.7 B). We modeled the eIF3b β-propeller domain with the help of the 40S•eIF1•eIF1A•eIF3 structure (Figures 2.10.7 A, C) as we detect only a part of it in the density (Figures 2.10.1 A - B). In py48S-open, weak densities of the PCI domains and eIF3b β-propeller appear only in low-resolution filtered map contoured at lower threshold (not shown) and were not modeled. It is not clear whether the much

weaker density for these regions of eIF3 in py48S-open is due to the lower quality of the data (fewer particles and lower resolution) or inherently greater flexibility or lower occupancy of these domains in py48S-open.

Remarkably, we see additional density for eIF3 in both complexes, at the subunit interface near h44, uS12, and eIF2γ (Figures 2.10.1 A - B). Based on its characteristic shape and dimensions, we assigned this density to the trimeric subcomplex composed of the β-propeller domain of eIF3i, ~30 residues from the eIF3b C-terminal domain (CTD) and ~50 residues from the N-terminal domain (NTD) of eIF3g (324) (Figure 6.1.2 D; Movie S1; See Supplemental Experimental Procedures, Section 6.3). eIF3i is positioned in the vicinity of eIF2γ, and the NTD of eIF3g may directly contact eIF2γ in the py48S complexes (Figures 2.10.7 C - D). eIF3i was earlier predicted to bind in two possible positions at the solvent-exposed surface of the 40S subunit, either above or below the β-propeller domain of eIF3b (324). Neither of these configurations is consistent with the density we observe, suggesting that eIF3 undergoes a significant rearrangement undetected by prior models, perhaps on binding mRNA. This position, in which eIF3i holds eIF3g against eIF2γ and by consequence promotes the intricate TC/eIF1A/eIF1 interaction network, might explain the suppression of eIF3i and eIF3g mutant phenotypes by overexpression of eIF1 or eIF1A and the formation of aberrant 43S complexes observed in the absence of these subunits (336, 337). This configuration also places eIF3i and eIF3g along the path of mRNA through the decoding center, consistent with the scanning defects observed for mutants of these subunits (336).

We also observe density in both complexes for a cluster of five α helices in a pocket formed by h11, h24, h27, h44, and uS15 that has been putatively assigned to a predicted helix-rich segment in the eIF3c-NTD (Figure 2.10.7 D; See Supplemental Experimental Procedures,

Section 6.3), which we connect to the eIF3c PCI domain by a ~30-residue flexible linker. The remaining N-terminal residues of eIF3c likely emanate from the five-helix cluster and mediate the known interaction of eIF3c-NTD with eIF1 (139, 205), which appears to enhance the stability of eIF1 within the PIC (146, 257). We therefore tentatively assigned the globular density in contact with eIF1 as the N-terminal region of the eIF3c-NTD (1–90 residues), with a single α-helix near h24 modeled in the density (Figure 2.10.7 D; see Supplemental Experimental Procedures, Section 6.3). This moiety approaches the surface of eIF1 identified as an eIF3c-binding surface (338).

Closer to the subunit interface, we detect density in both complexes for an extended helical region spanning h14, h44, and h27 (Figures 2.10.1, 6.1.2 E, and 6.1.7 D). These helices have been provisionally assigned to a region in the CTD of eIF3a (See Supplemental Experimental Procedures, Section 6.3) predicted to have long helices (339), and they bridge the β-propeller domain of eIF3i and the putative eIF3c-NTD moiety near eIF1. This assignment places the extreme C-terminal ~100 residues of eIF3a not modeled here in the vicinity of the TC, consistent with a known eIF3a-CTD interaction with eIF2 (145). We suggest that the unassigned central portion of eIF3a projects away from the eIF3a PCI domain near the exit channel on the 40S solvent face subunit and passes through the mRNA entry channel and across the intersubunit face, connecting with the extended helices assigned to the eIF3a-CTD (Figures 2.10.7 A and C). As these extended helices approach the eIF3c-NTD (Figure 2.10.7 D), it appears that eIF3a and eIF3c together encircle the PIC. This proposal is consistent with previous observations that regions of the eIF3a CTD interact with 40S components at the mRNA entry channel (189, 260) as well as structural models for the yeast 40S•eIF1•eIF3 complex (324) and the pm43S (109),

and chemical and enzymatic footprinting data (261), all of which place eIF3 components at both the exit and entry channels on the solvent side of the 40S subunit.

## 2.5 DISCUSSION

Our py48S-open and py48S-closed structures contain density for eIF1, eIF1A, all three subunits of eIF2 bound to Met-tRNA$_i$, mRNA, and various components of eIF3. Although both structures are at lower resolution than the previously reported py48S (110), this is likely to be the result of the small fraction of the particles in these classes. Despite this, the presence of additional factors here results in an overall stabilization and better local resolution for eIF1, eIF2α, eIF2γ, and tRNA$_i$ in py48S-closed (Table 6.1.1).

We observe eIF2β in py48S-open complex, where it connects eIF1 and eIF1A on the body with tRNA$_i$ on the 40S head. These bridging interactions should stabilize both TC and eIF1 binding in the scanning PIC prior to achieving a perfect AUG:anticodon duplex in the P site. Being unique to py48S-open, the eIF2β contacts with eIF1 and tRNA$_i$ should specifically stabilize the scanning complex. Consistent with this prediction, substitutions at both interfaces decreased the probability of continued scanning at near-cognate UUG start codons in yeast cells—presumably enabling rearrangement to the closed complex without a perfect start codon:anticodon duplex in the P site—thus establishing that eIF2β/eIF1 and eIF2β/tRNA$_i$ contacts in py48S-open promote initiation accuracy *in vivo*. The network of eIF2β interactions with eIF1/eIF1A/tRNA$_i$ should also impede mRNA insertion into the mRNA channel at the P site: eIF2β is likely repositioned to allow mRNA recruitment. Modeling either the conformation of eIF2β in py48S-closed, where it no longer contacts eIF1 and eIF1A

(Figure 6.1.7 A), or the distinct conformations observed in archaeal βγ complexes (286, 335) (Figure 6.1.6 D), into py48S-open reveals unfettered access to the mRNA channel, supporting the notion that transient repositioning of eIF2β would allow mRNA recruitment and that eIF2β serves as a barrier to mRNA release during scanning.

These structures show how eIF3 can interact with TC and eIF1 close to the P site at the inter-subunit interface even while the majority of its contacts map to the remaining solvent-exposed surfaces of the 40S subunit. Based on our modeling, eIF3 appears to connect the entry and exit channels on the solvent face of the 40S subunit to the center of action at the P site. None of the core subunits of eIF3 has previously been observed at the subunit interface, except for eIF3j (114, 324), which was excluded from our study. We note however that our complexes contain mRNA, whereas all previous PICs with eIF3 lacked mRNA (109, 114, 324). It is likely that the position of eIF3i observed earlier (114, 324) may represent its position prior to mRNA binding and that the presence of mRNA in our complexes may have led to the previously unobserved conformation of eIF3 at the inter-subunit interface. Interestingly, the positioning of the eIF3b-CTD/eIF3i/eIF3g-NTD module observed here, in proximity to eIF2γ, might hinder the insertion of mRNA into the mRNA-binding channel of the ribosome, making it likely that this conformation exists only after mRNA loading and suggesting that it might lock mRNA into the scanning complex. We propose that eIF3 undergoes a substantial conformational change upon mRNA binding, relocating both the eIF3b-CTD/eIF3i/eIF3g-NTD module and portions of eIF3a and eIF3c to enable their interactions with eIF2 and eIF1 in the decoding center and thereby facilitate key steps in scanning and start codon recognition. This rearrangement may signal the presence of mRNA within the PIC to other eIFs, notably eIF2.

Integrating the complete array of structures described in this report allows us to propose a detailed scheme for assembly of the 43S PIC, mRNA recruitment to this complex, and subsequent steps of scanning and start codon recognition in the 48S PIC. In the empty 40S subunit, the position of the head with respect to the body ensures that the latch is closed (127). Binding of eIF1 and eIF1A to assemble the 40S•eIF1•eIF1A complex (Figure 6.1.4 A) leads to an 8° rotation of the head compared to the empty 40S subunit in 80S ribosomes (92) that likely facilitates binding of TC in the $P_{OUT}$ state to form the 43S PIC (Figures 6.1.4 B and 6.1.7 B). Notably, the latch remains closed in both the 40S•eIF1•eIF1A and py43S complexes observed here. The 40S head is further rotated 5°–6° in the structures of py43S (without eIF3) or pm43S (without eIF1/eIF1A) (109) relative to 40S•eIF1•eIF1A (Figure 6.1.7 B), which may facilitate mRNA recruitment.

The py48S-open and py48S-closed structures (Figures 2.10.1 A - B) illuminate a series of rearrangements that enable the PIC to first bind and scan the mRNA and then halt upon recognition of the start codon. In py48S-open, the presence of eIF1, eIF1A, TC, and eIF3 provokes an upward movement of the head away from the body, opening the latch and widening the mRNA channel between the body and head, and opening the P site, which leads to diminished contacts with $tRNA_i$ relative to the closed state (Figure 6.1.7 B). We propose that this open conformation enables lateral insertion of the 5′ end of the mRNA—facilitated by the eIF4F complex bound at the cap—onto the 40S subunit (not shown). Once loaded onto the mRNA, py48S-open would be poised for scanning: the mRNA is held loosely in the channel; $tRNA_i$ is not fully engaged with the P site; the eIF1A-NTT is disordered. eIF2β interacts with $tRNA_i$, eIF1, and eIF1A to both stabilize TC binding and help hold the mRNA in the channel to promote processive scanning. The relocation of the eIF3b-CTD/eIF3i/eIF3g-NTD module near h44 on the

intersubunit face, where it interacts with eIF2γ, would promote the same ends. As the open complex scans the mRNA, eIF5-mediated GTP hydrolysis by eIF2 occurs, but $P_i$ release is blocked by the presence of eIF1, itself stabilized in the complex by its interaction with the NTD of eIF3c.

In py48S-closed, recognition of the start codon results in downward movement of the head, driven by a change in the pitch of h28 and changes in the orientation of eIF2β, closing the latch and fixing the mRNA in the channel to arrest scanning. Head closure also brings P-site elements in the 40S body into contact with the ASL, locking Met-$tRNA_i$ into the P site. Both this constriction of the entry channel and the enclosure of the P site around $tRNA_i$ (Figure 6.1.7 B) are supported by recent hydroxyl radical probing of yeast 48S complexes reconstituted with AUG versus AUC start codons (199). These and other rearrangements stabilize the $P_{IN}$ state of the closed complex (Figure 6.1.7 B): interactions between eIF2β and the $tRNA_i$ are remodeled; eIF2β exchanges its contacts with eIF1 and eIF1A for those with the 40S head; the eIF1A-NTT interacts with the AUG:anticodon duplex. Other rearrangements deform eIF1; different portions of $tRNA_i$ are brought into contact with eIF1, adjusting its position and promoting its eviction from the 40S subunit, which provokes $P_i$ release and commits the complex to subunit joining. $P_i$ release may also trigger detachment of eIF3b-CTD/eIF3i/eIF3g-NTD from the subunit interface, paving the way for release of eIF2, binding of eIF5B, and joining of the 60S subunit. Conformational changes within the TC in py48S-closed also bring the eIF2α-D1 loop in contact with the key −3 nt upstream of the start codon to regulate AUG selection (110).

In summary, the py48S-open and py48S-closed structures described here address long-standing questions about various aspects of initiation. Comparison of these structures reveal how the PIC in the open state may facilitate both loading of the mRNA and subsequent scanning, all

while holding both TC and mRNA in place for processive inspection of codons within the P site. Upon recognition of the start codon, the PIC closes, both locking mRNA and $tRNA_i$ within the P site and preparing eIF1 for its departure from the complex. Our structures also reveal how eIF3, bound at the 40S solvent face, may encircle the PIC, linking the mRNA entry and exit channels with the locus of action near the P site.

## 2.6 AUTHOR CONTRIBUTIONS

J.L.L. and T.H. made the samples, collected and analyzed the data, determined the structures and wrote a first draft of the manuscript. L.M. and A.T. performed the genetic experiments. C.E.A. provided advice on eIF3 purification, helped characterize recombinant eIF3, and helped write the manuscript. J.R.L., A.G.H., and V.R. supervised the work and helped to write the manuscript.

## 2.7 ACKNOWLEDGEMENTS

We are grateful to S. Chen, C.G. Savva, K.R. Vinothkumar, G. McMullan, and the staff of FEI for technical support with cryo-EM; T. Darling and J. Grimmett for help with computing; X.C. Bai and S.H.W. Scheres for advice with EM data processing; and G. Murshudov for help with refinement of the atomic coordinates. We also thank A. Kelley and I.S. Fernandez for providing reagents. J.L.L. and T.H. were, respectively, supported by postdoctoral fellowships from FEBS and EMBO. This work was funded by grants to from the UK Medical Research Council (MC_U105184332), Wellcome Trust Senior Investigator award (WT096570), the

Agouron Institute and the Jeantet Foundation to V.R.; from the NIH (GM62128) formerly to J.R.L.; the Human Frontiers in Science Program (RGP-0028/2009) to A.G.H., J.R.L., and V.R.; and by the Intramural Research Program of the NIH (A.G.H., J.R.L., and C.E.A.).

## 2.8 FOOTNOTES



Supplemental Information includes seven figures, one table, five movies, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.06.033 and in Appendix x.

## 2.9 ACCESSION NUMBERS

Maps have been deposited in the EMDB with accession codes EMD: EMD-3047, EMD-3048, EMD-3049, EMD-3050 for the 40S•eIF1•eIF1A, py48S-closed, py48S-open, and py43S, respectively. Atomic coordinates have been deposited in the PDB with accession codes PDB: 3JAM, 3JAP, 3JAQfor 40S•eIF1•eIF1A, py48S-closed, and py48S-open, respectively.

## 2.10 FIGURES AND TABLES

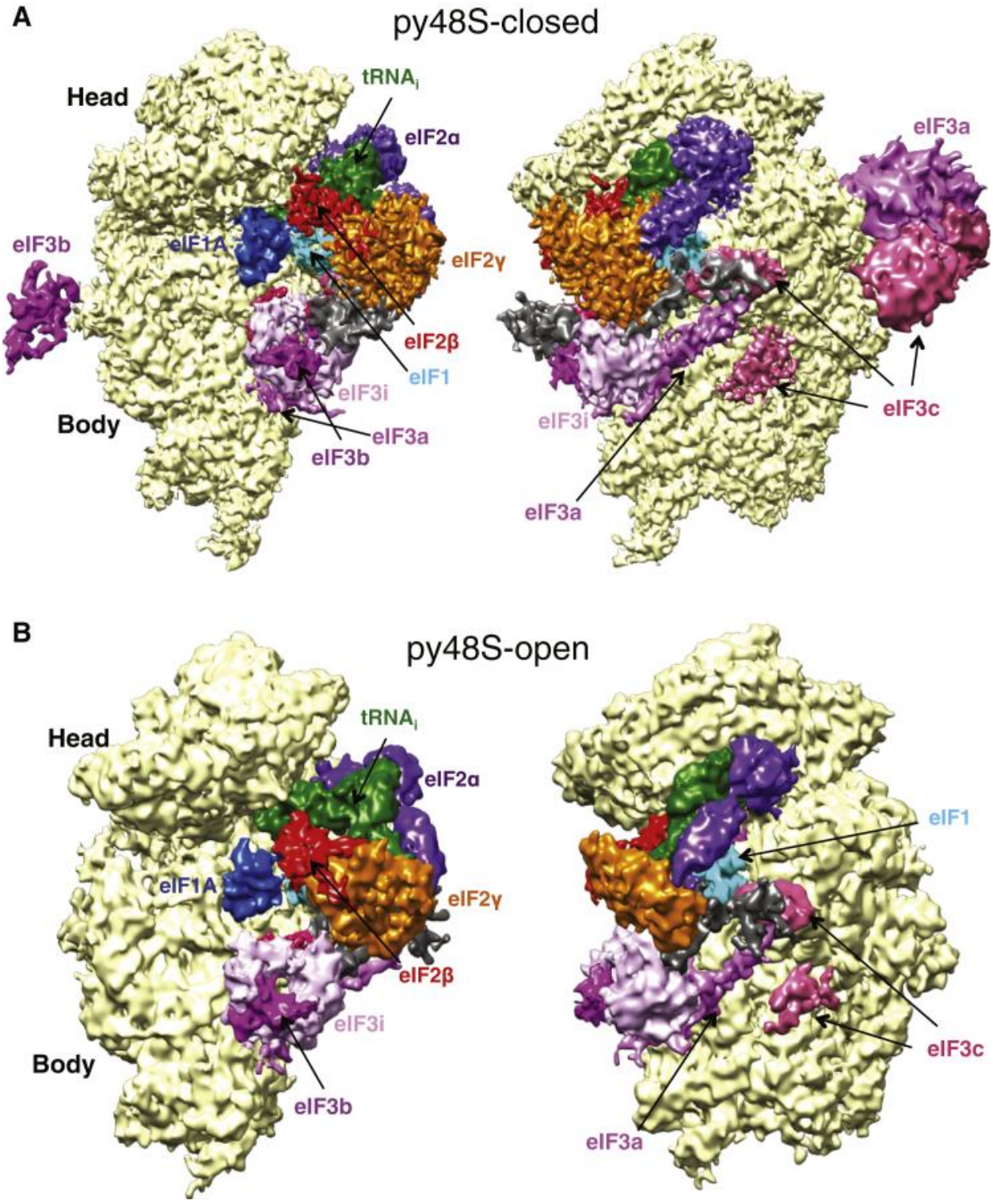


**Figure 2.10.1. Cryo-EM Maps of Eukaryotic 48S PICs**

(**A**) Two views of py48S-closed.

(**B**) Two views of py48S-open. Density for eIF3 is Gaussian-filtered. Unassigned density is in dark gray.

See also Figures 6.1.1-6.1.4; Table 6.1.1; and Movies S1 and S2.

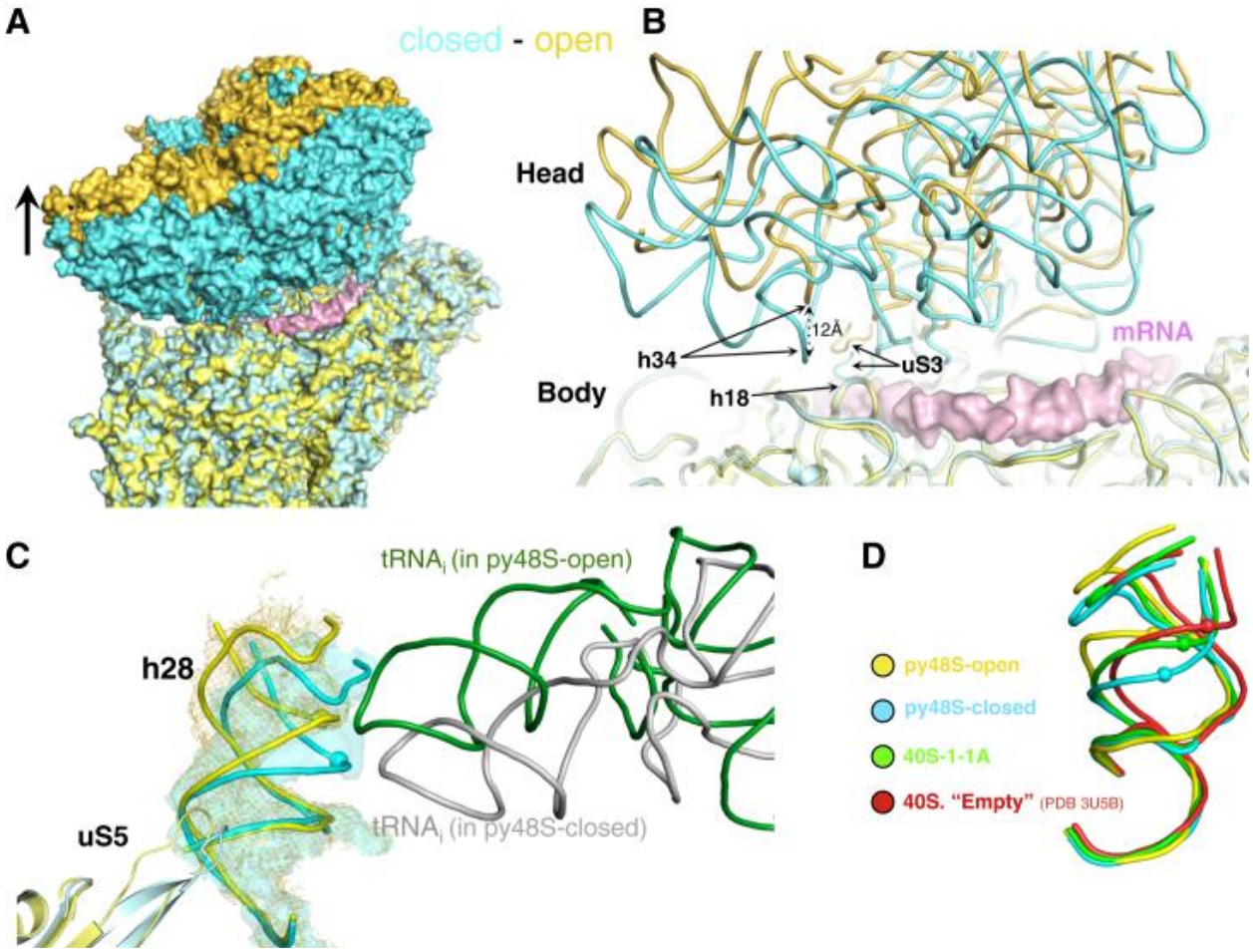

A
closed - open
B
Head
12Å
mRNA
h34
uS3
Body
h18
C
tRNAi (in py48S-open)
h28
uS5
tRNAi (in py48S-closed)
D
py48S-open
py48S-closed
40S-1-1A
40S. "Empty" (PDB 3U5B)

**Figure 2.10.2. Distinct Position of the 40S Head Widens the mRNA Entry Channel and Opens the Latch in the Open Complex**

(**A**) Front view of superposition of py48S-open (yellow) and py48S-closed (cyan), showing mRNA (pink) from py48S (Hussain et al., 2014) to highlight the complete mRNA channel.

(**B**) Superposition of refined models of py48S-open (yellow) and py48S-closed (cyan), indicating elements forming the latch.

(**C**) h28 in py48S-open (yellow) and py48S-closed (cyan) based on superposition of the two complexes, viewed from the A site. $tRNA_i$ and uS5 for the two complexes are also shown.

(**D**) h28 in empty 40S (red, PDB:3U5B), 40S•eIF1•eIF1A (green), py48S-open (yellow), and py48S-closed (cyan) complexes. Equivalent atoms in h28 are shown as spheres.

See also Figures 6.1.2 and 6.1.4 and Movies S3 and S4.

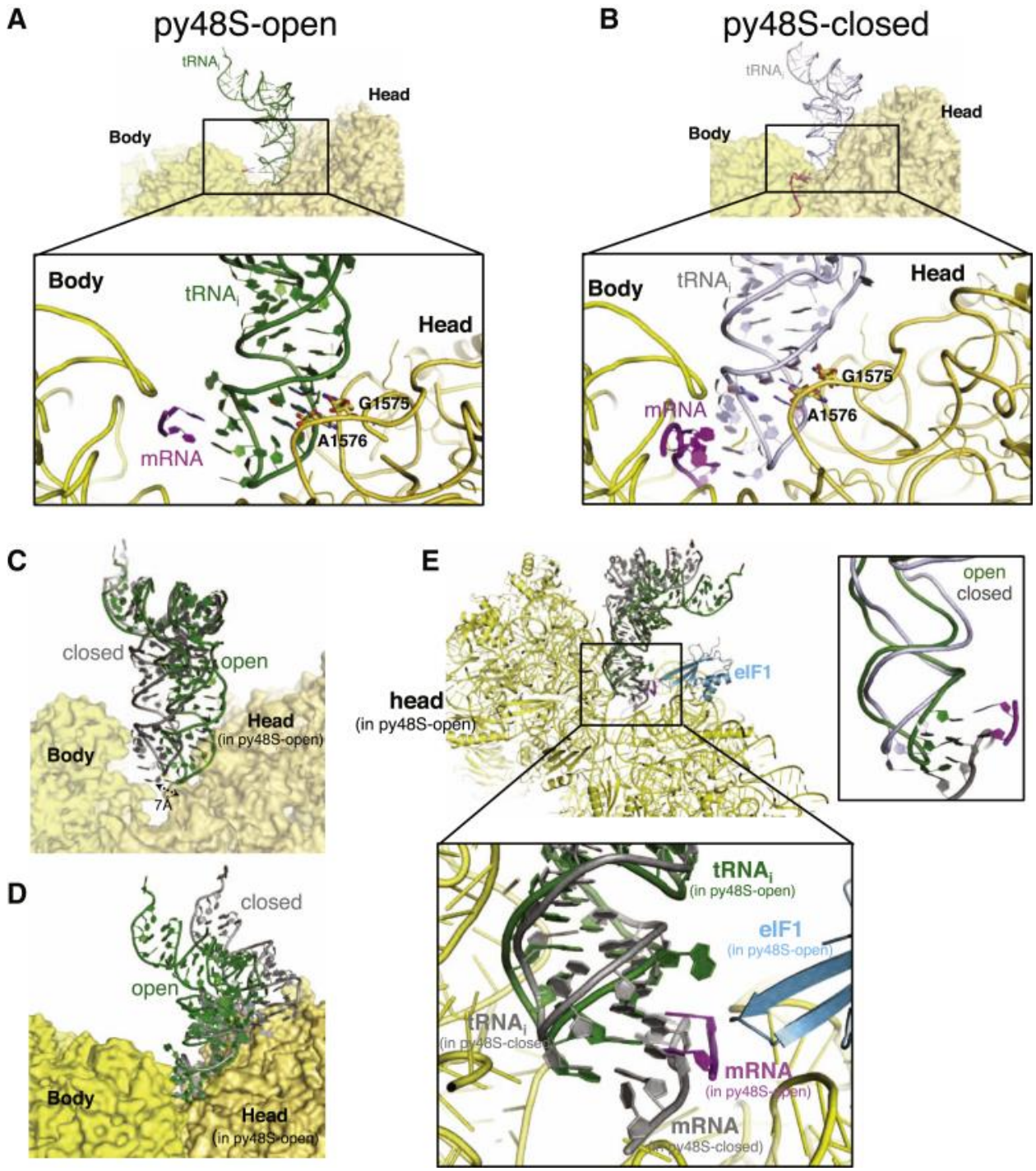

A
py48S-open
tRNAi
Head
Body
Body
tRNAi
Head
G1575
A1576
mRNA
B
py48S-closed
tRNAi
Head
Body
Body
tRNAi
Head
G1575
A1576
mRNA
C
closed
open
Head
(in py48S-open)
Body
7Å
D
closed
open
Body
Head
(in py48S-open)
E
head
(in py48S-open)
eIF1
open
closed
tRNAi
(in py48S-open)
eIF1
(in py48S-open)
tRNAi
(in py48S-closed)
mRNA
(in py48S-open)
mRNA
(in py48S-closed)

**Figure 2.10.3. $tRNA_i$ Is Not Engaged with P-Site Elements of the 40S Body in the Open Complex**

(**A**) $tRNA_i$ in py48S-open, viewed from E site. The body and head of 40S are shown in lighter and darker shades of yellow. The zoomed view shows mRNA at the P site and recognition of conserved GC base pairs in ASL by rRNA bases. For clarity, 40S proteins and other factors are not shown.

(**B**) The $tRNA_i$ in py48S-closed viewed as in (A).

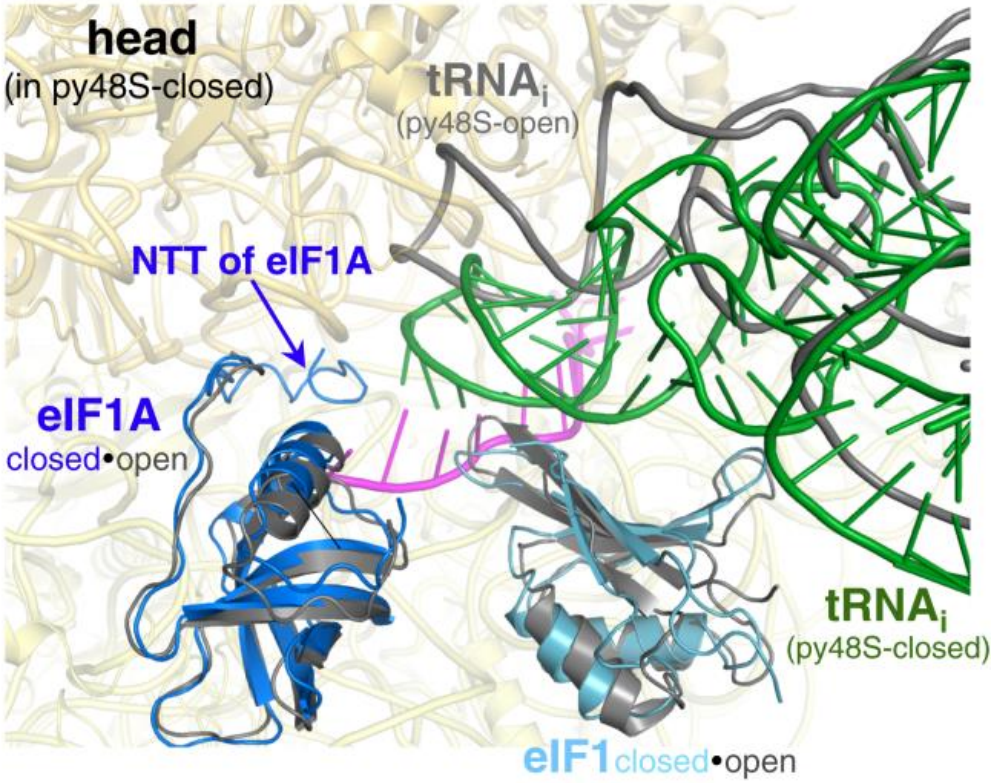
(**C**) Superposition of the 40S body reveals distinct locations of $tRNA_i$ in the P site of py48S-open (green) and py48S-closed (gray). The body and head of py48S-open complex are shown. The two ASLs are separated by about 7 Å in the P site.

(**D**) Superposition of the 40S head of py48S-open (green) and py48S-closed (gray).

(**E**) Superposition of two complexes as in (D), viewed from the A site. The mRNA of py48S-closed is in gray. Inset shows the superposition of the two positions of $tRNA_i$ and interacting mRNA codon.

See also Figure 6.1.5.

A


head
(in py48S-closed)
tRNAi
(py48S-open)
NTT of eIF1A
eIF1A
closed•open
tRNAi
(py48S-closed)
eIF1 closed•open

B

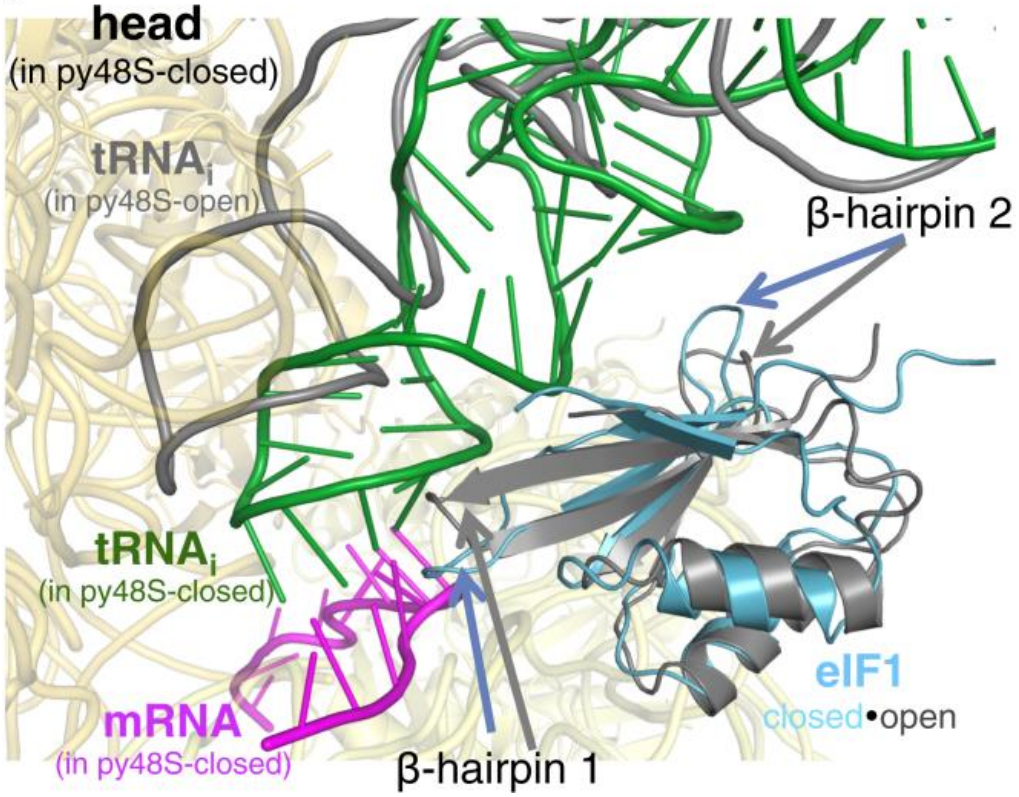

head
(in py48S-closed)
tRNAi
(in py48S-open)
β-hairpin 2
tRNAi
(in py48S-closed)
eIF1
closed•open
mRNA
(in py48S-closed)
β-hairpin 1

**Figure 2.10.4. Contacts of eIF1A-NTT and eIF1 with $tRNA_i$ Restricted to the Closed Complex**

(**A**) Superposition of the open and closed complexes with the ligands of py48S-closed shown in color while those of py48S-open are in gray. Only the 40S of py48S-closed is shown (yellow). The zoomed view shows the NTT of eIF1A in the two complexes.

(**B**) Superposition of the open and closed complexes as in (A).

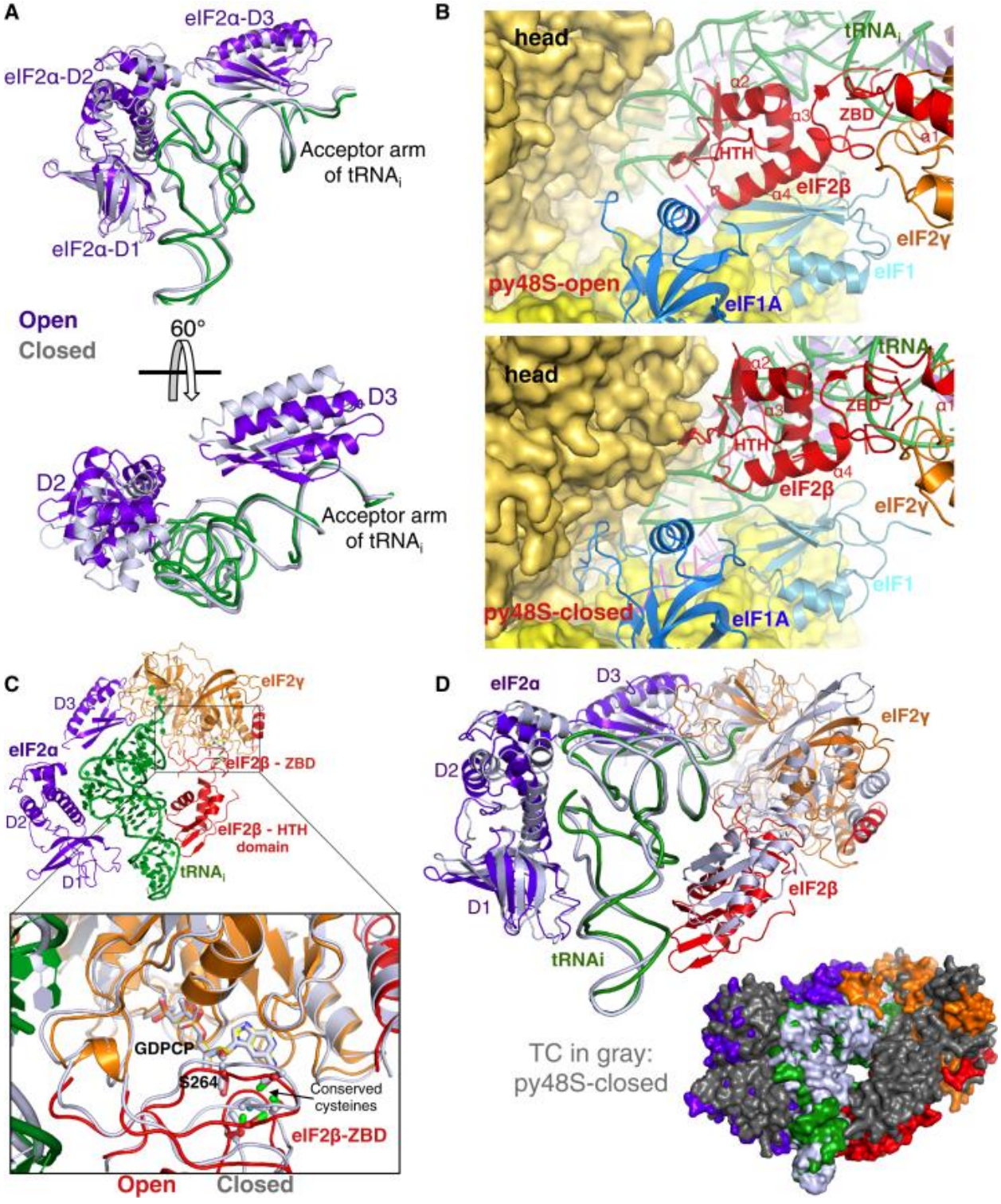

A
eIF2α-D3
eIF2α-D2
Acceptor arm of tRNAi
eIF2α-D1
Open
Closed
60°
D3
D2
Acceptor arm of tRNAi
B
head
tRNAi
α2
α3
ZBD
α1
HTH
α4
eIF2β
eIF2γ
eIF1
py48S-open
eIF1A
head
tRNAi
α2
α3
ZBD
α1
HTH
α4
eIF2β
eIF2γ
eIF1
py48S-closed
eIF1A
C
eIF2γ
D3
eIF2α
eIF2β - ZBD
D2
eIF2β - HTH domain
tRNAi
D1
GDPCP
S264
Conserved cysteines
eIF2β-ZBD
Open
Closed
D
eIF2α
D3
eIF2γ
D2
eIF2β
D1
tRNAi
TC in gray:
py48S-closed

**Figure 2.10.5. Distinctive Interactions of eIF2β with eIF1, eIF1A, and $tRNA_i$ Occlude the mRNA Channel in py48S-Open**

(**A**) Conformational changes in eIF2α based on superposition of the TC coordinates using $tRNA_i$ as the reference. The eIF2α and $tRNA_i$ of py48S-open are shown in color and those of py48S-closed are in gray.

(**B**) Position of eIF2β with respect to $tRNA_i$, eIF1, eIF1A, and 40S head in py48S-open and py48S-closed.

(**C**) Similar position for the ZBD of eIF2β in both complexes with respect to eIF2γ. The eIF2β, eIF2γ and $tRNA_i$ of py48S-open are shown in color while those of py48S-closed are in gray. Ser264 is shown as sticks near conserved cysteines.

(**D**) Cartoon and surface representations of the superimposition of TC coordinates in py48S-open (color) and closed (gray) complexes based on $tRNA_i$ as reference. It shows the internal conformational change within TC during transition from the open to the closed conformation. While D2 and the helix connecting the D1 and the D2 domains of eIF2α experience an internal rearrangement, eIF2α-D3, eIF2γ and eIF2β rotate together around the acceptor arm of the tRNA. See also Figures 6.1.6 and 6.1.7 and Movie S5.

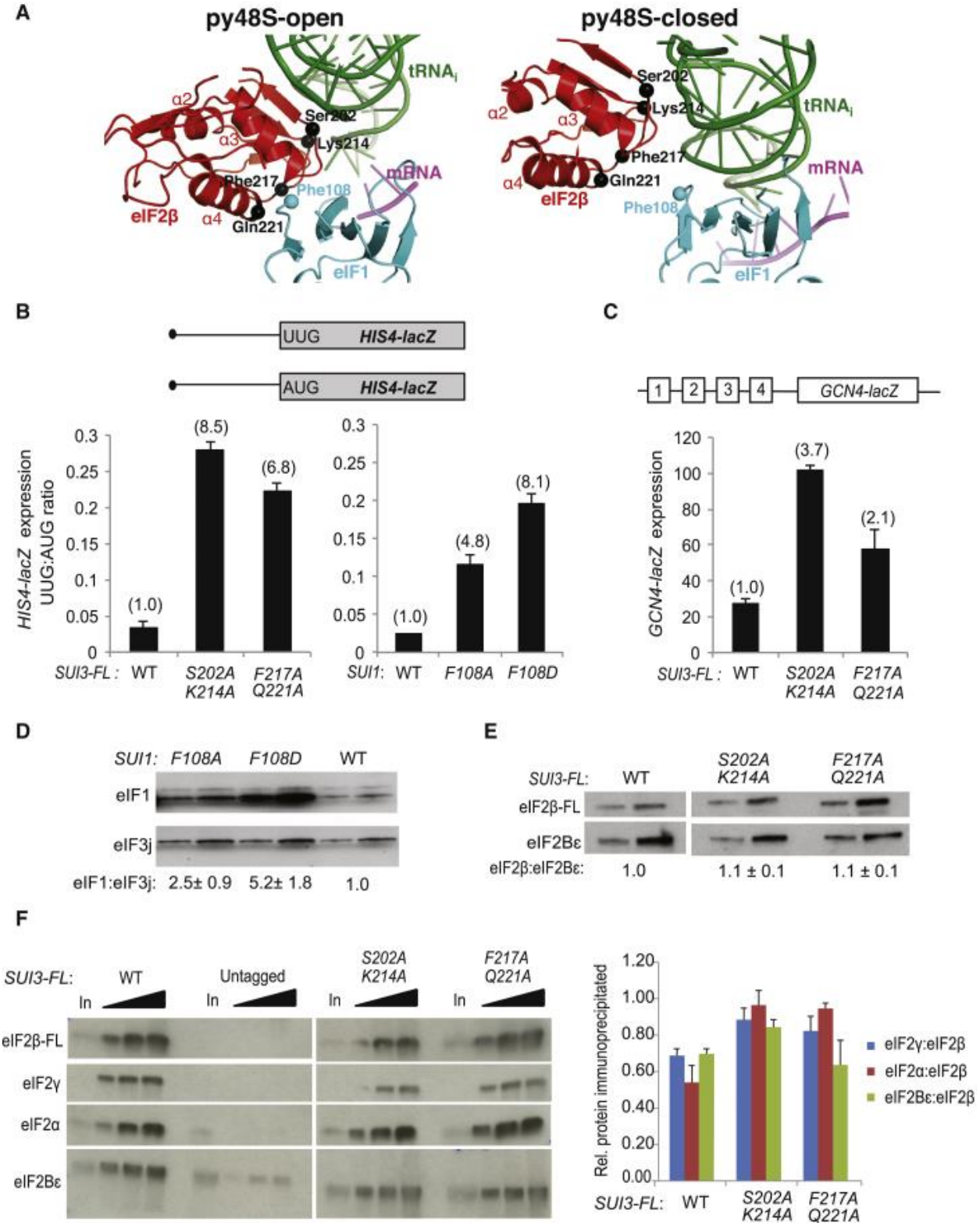

A
py48S-open
py48S-closed
tRNAi
α2
α3
α4
Ser202
Lys214
Phe217
Gln221
mRNA
Phe108
eIF2β
eIF1
B
UUG HIS4-lacZ
AUG HIS4-lacZ
HIS4-lacZ expression UUG:AUG ratio
(8.5)
(6.8)
(1.0)
(4.8)
(8.1)
SUI3-FL : WT S202A K214A F217A Q221A
SUI1: WT F108A F108D
C
1 2 3 4 GCN4-lacZ
GCN4-lacZ expression
(3.7)
(2.1)
(1.0)
D
SUI1: F108A F108D WT
eIF1
eIF3j
eIF1:eIF3j: 2.5± 0.9 5.2± 1.8 1.0
E
SUI3-FL: WT S202A K214A F217A Q221A
eIF2β-FL
eIF2Bε
eIF2β:eIF2Bε: 1.0 1.1 ± 0.1 1.1 ± 0.1
F
SUI3-FL: WT Untagged S202A K214A F217A Q221A
In
eIF2β-FL
eIF2γ
eIF2α
eIF2Bε
Rel. protein immunoprecipitated
eIF2γ:eIF2β
eIF2α:eIF2β
eIF2Bε:eIF2β

**Figure 2.10.6. Genetic Evidence that eIF2β Interactions with the $tRNA_i$ ASL and eIF1 Preferentially Stabilize py48S-Open to Impede Initiation at Near-Cognate UUG Codons *In Vivo***

(**A**) Positions of eIF1, eIF2β, and $tRNA_i$ in the py48S-open and py48S-closed, with residues substituted in genetic studies shown as spheres.

(**B**) Expression of *HIS4-lacZ* reporters with AUG or UUG start codons in strains of the indicated *SUI3* or *SUI1* genotypes, expressed as mean (± SEM) ratios of UUG- to AUG-reporter expression with fold-changes relative to WT in parentheses.

(**C**) Expression of the *GCN4-lacZ* reporter expressed as mean (± SEM) units of β-galactosidase.

(**D** and **E**) Western analysis of eIF1 (D) or eIF2β (E) proteins in whole-cell extracts (WCEs), with eIF3j or eIF2Bε analyzed as loading controls, reported as mean (± SEM) eIF1:eIF3j ratios or eIF2β:eIF2Bε ratios, normalized to the WT ratios, determined from biological replicates. Lanes have been cropped from the same gels.

(F) WCEs were immunoprecipitated with FLAG affinity resin and immune complexes subjected to Western analysis to detect Flag-eIF2β and co-immunoprecipitated eIF2γ, eIF2α and eIF2Bε, resolving 1x, 2x, or 3x amounts in successive lanes. In, 20% input WCEs. Western signals were quantified to yield mean (± SEM) recoveries of eIF2γ, eIF2α, or eIF2Bε normalized to Flag-eIF2β. Lanes have been cropped from the same gels. Error bars represent the SEM from three biological replicates.

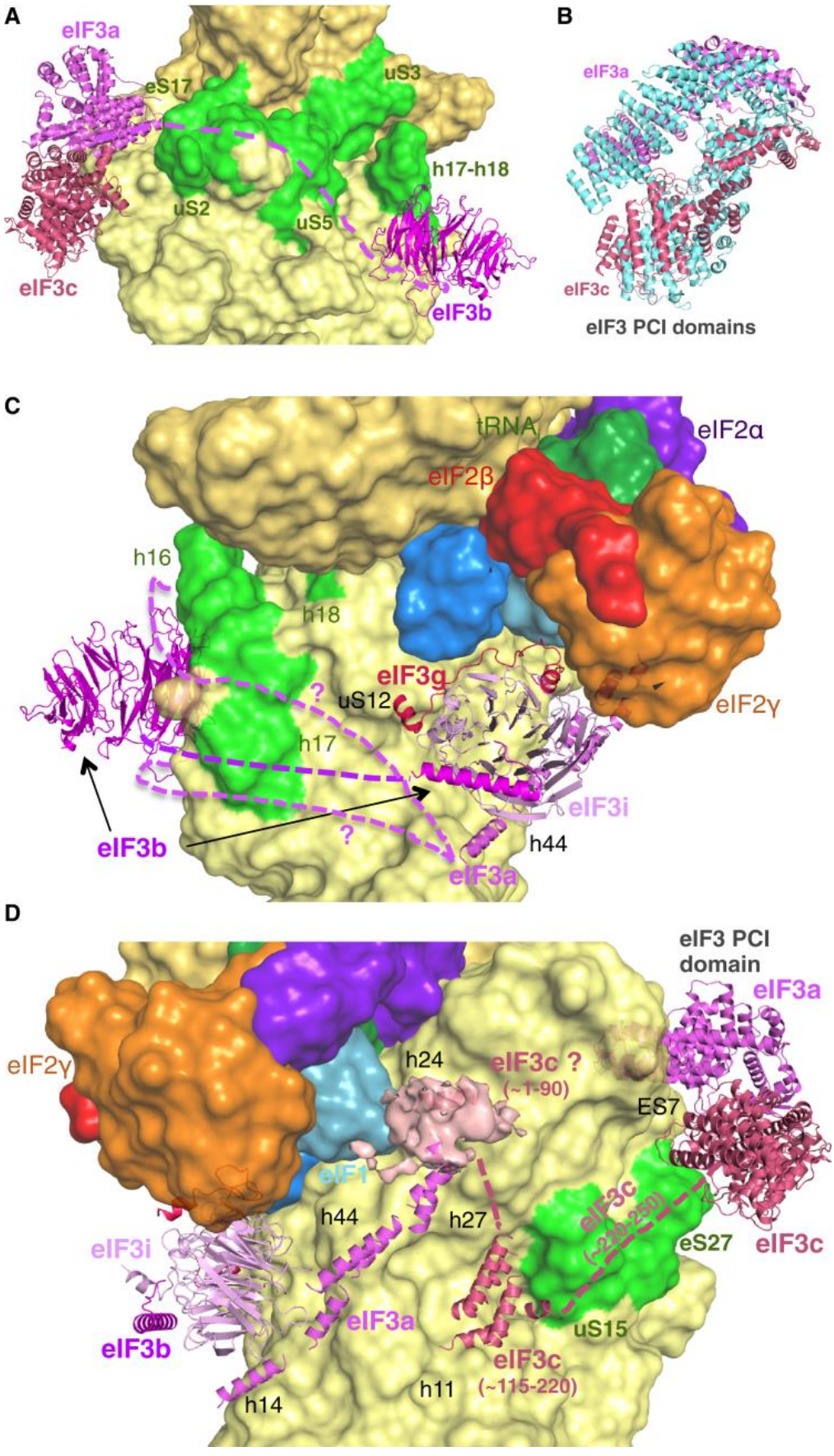

A
eIF3a
eS17
uS3
h17-h18
uS2
uS5
eIF3c
eIF3b
B
eIF3a
eIF3c
eIF3 PCI domains
C
tRNA
eIF2α
eIF2β
h16
h18
eIF3g
uS12
eIF2γ
h17
eIF3i
eIF3b
h44
eIF3a
D
eIF3 PCI domain
eIF3a
eIF2γ
h24
eIF3c ?
(~1-90)
ES7
eIF1
h44
h27
eIF3c
(~220-230)
eS27
eIF3c
eIF3i
eIF3b
eIF3a
uS15
eIF3c
(~115-220)
h11
h14

**Figure 2.10.7. Structural Arrangement of eIF3 Components in 48S PICs**

(**A**) Locations of the eIF3a/eIF3c PCI domains and β-propeller of eIF3b at different positions on the solvent-exposed surface of the 40S, highlighting rRNA helices and ribosomal proteins (green) predicted to bind to eIF3a. The proposed path of the unassigned central portion of the eIF3a-CTD connecting the PCI domain to the subunit interface is shown as a dashed purple line.
(**B**) Lateral displacement of eIF3a/eIF3c PCI domains in py48S-closed versus their positions in yeast 40S•eIF1•eIF1A•eIF3 (PDB: 4UER).
(**C**) Trimeric eIF3b-CTD/eIF3i/eIF3g-NTD subcomplex is shown near h44 and interacting with eIF2γ and the 40S interface surface. The β-propeller of eIF3b is also shown. Two alternative proposed paths of the eIF3a-CTD connecting the PCI domain to the bundle of helices below the eIF3i β-propeller are shown as dashed purple lines.
(**D**) A cluster of helices tentatively assigned to eIF3c is located near h11 and uS15 (green). A globular density with a single modeled helix is tentatively assigned to the eIF3c-NTD in proximity to eIF1 and h24. The proposed path of a linker connecting the cluster of helices to the eIF3c PCI domain is shown as a dashed magenta line. Long helices tentatively assigned to eIF3a bridge the eIF3i β-propeller and h44 with the putative eIF3c-NTD and h24.

# 3. A Network of eIF2β Interactions with eIF1 and Met-tRNA$_i$ Promotes Accurate Start Codon Selection by the Translation Initiation Complex

Adapted and reprinted with permission from:

A Network of eIF2β interactions with eIF1 and Met-tRNA$_i$ promotes accurate start codon selection by the translation initiation complex

Laura Marler, Anil Thakur, Alan G. Hinnebusch

Nucleic Acids Research 2018

# A network of eIF2β interactions with eIF1 and Met-tRNA$_i$ promotes accurate start codon selection by the translation preinitiation complex

Anil Thakur[*], Laura Marler[*], and Alan G. Hinnebusch[†]

Division of Molecular and Cellular Biology, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, Bethesda, Maryland 20892

[*]The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint first authors.

[†]To whom correspondence should be addressed. Tel: +1-301-496-4480; Email: ahinnebusch@nih.gov

Keywords: translation, initiation, regulation, eIF2β, eIF1, Met-tRNA$_i$, ribosome, yeast

Running title: eIF2β/eIF1/Met-tRNA$_i$ interactions regulate AUG selection

## 3.1 ABSTRACT

In translation initiation, a 43S preinitiation complex (PIC) containing eIF1 and a ternary complex (TC) of GTP-bound eIF2 and Met-RNA$_i$ scans the mRNA for the start codon. AUG recognition triggers eIF1 release and rearrangement from an open PIC conformation to a closed state with more tightly-bound Met-tRNA$_i$ ($P_{IN}$ state). Cryo-EM models reveal eIF2β contacts with eIF1 and Met-tRNA$_i$ exclusive to the open complex that should destabilize the closed state. eIF2β or eIF1 substitutions disrupting these contacts increase initiation at UUG codons, and compound substitutions also derepress translation of *GCN4*, indicating slower TC recruitment. The latter substitutions slow TC loading while stabilizing TC binding at UUG codons in reconstituted PICs, indicating a destabilized open complex and shift to the closed/$P_{IN}$ state. An eIF1 substitution that should strengthen the eIF2β:eIF1 interface has the opposite genetic and biochemical phenotypes. eIF2β is also predicted to restrict Met-tRNA$_i$ movement into the closed/$P_{IN}$ state, and substitutions that should diminish this clash increase UUG initiation *in vivo* and stabilize Met-tRNA$_i$ binding at UUG codons *in vitro* with little effect on TC loading. Thus, eIF2β anchors eIF1 and TC to the open complex, enhancing PIC assembly and scanning, while impeding rearrangement to the closed conformation at non-AUG codons.

## 3.2 INTRODUCTION

The process of translation initiation selects the start codon for protein synthesis, determining the reading frame for elongation and N-terminus of the polypeptide, and also influencing the rate at which the protein is synthesized. Accordingly, accurate identification of the translation start codon is critical to ensure the synthesis of the correct cellular proteins in the proper amounts. In eukaryotes, translation initiation occurs via a scanning mechanism, wherein the small (40S) subunit of the ribosome recruits methionyl initiator tRNA (Met-tRNA$_i$) in a ternary complex (TC) with GTP-bound eukaryotic initiation factor 2 (eIF2), in a reaction stimulated by factors eIF1, eIF1A, and eIF3. The resulting 43S preinitiation complex (PIC) attaches to the 5' end of mRNA and scans the mRNA leader for an AUG start codon. The nucleotide sequence immediately surrounding the start codon—the AUG context—particularly at the -3 and +4 positions (numbered from the A of AUG (+1)) also influences the efficiency of start codon selection. In the scanning PIC, eIF1 and eIF1A promote an open, scanning-conducive conformation of the 40S subunit with TC bound in a relatively unstable conformation, "P$_{OUT}$", which facilitates the inspection of successive triplets in the peptidyl (P) decoding site for complementarity with the anticodon of Met-tRNA$_i$. The GTP bound to eIF2 can be hydrolyzed, stimulated by GTPase activating protein eIF5, but eIF1 blocks release of inorganic phosphate (P$_i$) at non-AUG codons. Start codon recognition triggers dissociation of eIF1 from the 40S subunit, enabling both P$_i$ release from eIF2-GDP·P$_i$ and more stable TC binding to the PIC, with Met-tRNA$_i$ more fully accommodated in the "P$_{IN}$" state. Subsequent dissociation of eIF2-GDP and other eIFs from the 48S PIC enables eIF5B-catalyzed subunit joining and formation of an 80S initiation complex ready to commence protein synthesis (12).

eIF1 plays a dual role in translation initiation. It promotes the open conformation of the PIC, to which TC rapidly loads in the $P_{OUT}$ conformation, and it ensures initiation fidelity by blocking $P_i$ release and impeding isomerization to the closed/$P_{IN}$ state at non-AUG codons or start codons in poor context (Fig. 3.10.1 A). Structural analyses of the PIC reveal that eIF1 and eIF1A promote rotation of the 40S head relative to the body, which likely enhances TC binding, while eIF1 clashes with Met-$tRNA_i$ in the $P_{IN}$ state. Hence, eIF1 dissociation from the 40S subunit is required for start codon recognition, and mutations that weaken eIF1 binding to 40S subunits confer dual defects *in vivo*: (i) they reduce the rate of TC loading, since eIF1 promotes TC binding to the open conformation of the PIC; and (ii) they increase initiation at near cognate codons or AUG codons in poor context, by destabilizing the open/$P_{OUT}$ state and favoring rearrangement to the closed/$P_{IN}$ state (12, 340). A reduced rate of TC loading resulting from such eIF1 mutations confers derepressed translation of *GCN4* mRNA *in vivo* (the $Gcd^-$ phenotype), as slower TC binding to PICs scanning the *GCN4* mRNA leader allows inhibitory upstream open reading frames (uORFs) to be bypassed in favor of reinitiation further downstream at the *GCN4* coding sequence (295). Increased initiation at near-cognate codons in such eIF1 mutants also restores translation of *his4-303* mRNA, lacking the AUG start codon, by elevating initiation at an in-frame UUG triplet at the third codon (the $Sui^-$ phenotype) (341) (Fig. 3.10.1 B). The AUG codon of the eIF1 gene itself (*SUI1* in yeast) occurs in suboptimal context and the frequency of its recognition is inversely related to eIF1 abundance, establishing a negative feedback loop that maintains proper eIF1 levels (212, 215). Whereas overexpressing WT eIF1 suppresses initiation at its own suboptimal AUG codon, eIF1 mutants defective for 40S binding relax discrimination against poor context and

increase the translational efficiency of *SUI1* mRNA, elevating expression of such eIF1 variants. These effects have been attributed to the altered rates of eIF1 dissociation from the scanning PIC, impeding or enhancing, respectively, inappropriate isomerization to the closed state at non-AUG codons or at AUG codons in poor context (214, 342).

Many mechanistic aspects of the translation initiation process remain unclear, including the exact molecular mechanism by which recognition of a start codon triggers isomerization of the PIC from the open/$P_{OUT}$ state to the closed/$P_{IN}$ state and the cessation of scanning. Cryo-electron microscopy (cryo-EM) structures of two distinct py48S complexes, both containing eIF1, eIF1A, mRNA and TC, but appearing to represent distinct intermediates in scanning and start codon recognition (py48S-open and py48S-closed), shed light on this process (111). The py48S-open complex exhibits an upward movement of the 40S head from the body that widens both the mRNA binding cleft and the P site, eliminating certain 40S contacts with the mRNA and Met-$tRNA_i$ that occur only in py48S-closed. Conversely, the closed structure shows a constricted mRNA channel and narrowed P site that encloses Met-$tRNA_i$. Comparing these two structures enables predictions about the factors and specific residues preferentially stabilizing either the open or the closed state of the PIC. Further, these structures show clear density for the globular portion of the β-subunit of eIF2, unresolved in previous PIC structures, allowing predictions about the molecular functions of eIF2β in TC loading and start codon recognition.

Evidence from the cryo-EM structures indicates that the helix-turn-helix (HTH) domain of eIF2β forms contacts with eIF1 that are specific to the open conformation of the PIC, as the HTH domain moves away from eIF1 during rearrangement to the closed complex (111) (Fig. 3.10.2 A). We reasoned that substitutions in the eIF2β HTH domain that perturb these contacts would promote premature eIF1 release, and thereby increase the likelihood of isomerization to the closed state at UUG codons or AUG codons in suboptimal context, conferring the $Sui^-$ phenotype and increasing eIF1 expression from *SUI1* mRNA containing native, poor context. The loss of contacts between eIF1 and eIF2β specific to the open state is also expected to reduce eIF1 occupancy of the open PIC and reduce the rate of TC loading to confer the $Gcd^-$ phenotype. This is the same collection of phenotypes noted above for eIF1 mutations that weaken its 40S binding.

The HTH domain of eIF2β also forms distinct contacts with the Met-$tRNA_i$ in the py48S-open and py48S-closed structures. Downward movement of the 40S head towards the body in the open-to-closed transition is associated with movement of the eIF2β-HTH away from the anticodon stem-loop (ASL) of Met-$tRNA_i$ (111) (Fig. 7.2.1 A-B). We hypothesized that eIF2β-HTH: Met-$tRNA_i$-ASL interactions specific to the open state help to stabilize the Met-$tRNA_i$ in the open conformation in the absence of a perfect codon-anticodon duplex and other canonical P site tRNA contacts restricted to the closed complex. They might also enhance the rate of TC loading to the open conformation. If so, then eIF2β substitutions that perturb these eIF2β:Met-$tRNA_i$ contacts should destabilize the open PIC conformation and confer the dual $Sui^-$ and $Gcd^-$ phenotypes predicted above for disruptions at the eIF1:eIF2β interface.

In addition to contacting the Met-tRNA$_i$-ASL in the open conformation, the eIF2β-HTH is predicted to clash with the D-loop of Met-tRNA$_i$ bound in the P$_{IN}$ state when eIF2β is bound in the open conformation, as revealed by an overlay of the py48S-open and py48S-closed structures (Fig. 7.2.1 C). This clash would occur as a result of Met-tRNA$_i$ movement relative to eIF2β in the downward movement of the 40S head toward the body during isomerization to the closed state and is avoided by displacement of eIF2β in py48S-closed relative to py48S-open (111). This structural evidence suggests that eIF2β might act to sterically oppose transition of the PIC to the closed state before an appropriate start codon is recognized by impeding accommodation of Met-tRNA$_i$ in the P$_{IN}$ state. Hence, we predicted that substitutions in the eIF2β HTH domain that diminish the predicted clash would favor isomerization to the closed state and increase its likelihood at non-AUG or suboptimal AUG codons. However, these substitutions are not predicted to destabilize the open complex, and should therefore have a minimal effect on the rate of TC loading. As such, they should confer Sui$^-$ but not Gcd$^-$ phenotypes. This phenotypic signature is rare among known Sui$^-$ mutations, most of which act through destabilization of the open PIC, but it has been observed in Sui$^-$ substitutions of eIF1 Loop 2, which act by a similar mechanism to alleviate the clash between eIF1 in the open conformation and Met-tRNA$_i$ in the closed/P$_{IN}$ state (200).

In an effort to support the physiological relevance of the py48S-open and py48S-closed structures reported by Llacer et al., we previously generated mutations in eIF2β and eIF1 at the eIF1:eIF2β interface and found that they reduced stringent selection of AUG start codons in the manner expected from a role for this interface in promoting the open

conformation of the PIC (111). Here, we expanded on these genetic findings with analysis of additional substitutions perturbing the eIF1:eIF2β interface in the open complex, and observed opposing effects on initiation fidelity for substitutions that either weaken this mutual interface (promoting eIF1 release and reducing the stringency of start codon selection) or strengthen it (stabilizing eIF1 binding and increasing the requirement for an AUG start codon in preferred context). We further demonstrated that eIF1:eIF2β interactions promote recognition of optimal Kozak context in addition to stringent AUG selection, as well as the rapid recruitment of TC to the open complex *in vivo*. Importantly, we recapitulated the effects of exemplar substitutions that weaken eIF1:eIF2β interactions in reducing TC loading and increasing near-cognate start codon selection in a fully purified translation initiation system. A similar *in vitro* reconstitution was achieved for substitutions at the eIF2β:Met-tRNA$_i$ interface found exclusively in the open conformation that both impair TC loading and stringent AUG selection *in vivo*. Finally, we provide genetic and biochemical evidence that a predicted clash between eIF2β in the open complex with Met-tRNA$_i$ in the closed state helps to restrict transition to the closed conformation to AUG start codons, without perturbing the open conformation of the PIC. Together, these findings firmly establish three critical roles for eIF2β in accurate start codon recognition: (i) eIF2β is involved in anchoring eIF1 specifically to the open PIC, promoting rapid TC loading and helping to prevent premature release of eIF1 at non-AUG codons; (ii) eIF2β stabilizes binding of Met-tRNA$_i$ specifically to the open conformation of the PIC, when many contacts between Met-tRNA$_i$ and the 40S subunit are not yet formed; and (iii) eIF2β functions similarly to eIF1 by impeding rearrangement of Met-tRNA$_i$ to the P$_{IN}$ state via a steric clash with Met-tRNA$_i$ D-loop, thus helping to limit this transition to AUG start

codons *in vivo*. Hence, eIF2β emerges as a key component of the complex network of physical interactions within the PIC that promote ribosomal scanning and restrict initiation to AUG codons in optimum context.

## 3.3 MATERIALS AND METHODS

### 3.3.1 Plasmid and yeast strain constructions

Yeast strains used in this study are listed in supplementary Table 7.1.2. Strains ATY100 and ATY122 to ATY128 harboring mutant *SUI1* alleles on single copy (sc) *LEU2* plasmids were derived from strain JCY03 (*MATa ura3-52 leu2-3, leu2-112 trp1Δ-63 his4-301(ACG) sui1Δ::hisG* p1200 (sc *URA3 SUI1*) by plasmid shuffling, using growth on 5-fluoorotic acid (5-FOA) medium to select for loss of *URA3* plasmid p1200 containing WT *SUI1* (plasmid-shuffling). The QuikChange site-directed mutagenesis system (Stratagene) was employed with the primers indicated in supplemental Table 7.1.3 to generate all mutant *SUI1* plasmids listed in supplemental Table 7.1.1, using as templates plasmids pJCB101 for yeast plasmids, or pTYB2-eIF1 for plasmids used to express recombinant eIF1 proteins in bacteria for purification. The high-copy (hc) plasmid p4385 (hc TC, *TRP1*) was generated from p1780-IMT (hc TC, *URA3*) by the "marker swap" approach using the *URA3*-to-*TRP1* converter plasmid pUT11 (343).

To generate strains LMY103-LMY117, LMY130-LMY134, and LMY142, expressing eIF2β variants, strain KAY18 was transformed to $Leu^+$ with low-copy (lc) *LEU2* plasmids harboring WT or mutant *SUI3* alleles constructed from plasmid p920

(Table 7.1.1) using the appropriate primers (Table 7.1.3) and the QuikChange site-directed mutagenesis system (Stratagene) according to the manufacturer's directions. The S202A/K214A double substitution was made sequentially by introducing the S202A mutation into a plasmid containing K214A in a second round of site-directed mutagenesis. The resident WT *SUI3*, *URA3* plasmid p921 was evicted by counter-selection on 5-FOA medium. Strains LMY128, LMY129, and LMY137, used for purification of eIF2 variants containing eIF2β-S202A/K214A, -F217A/Q221A, or -E189R, respectively, were constructed from H3840 by plasmid shuffling to replace pAV1089 (containing *URA3* and WT *SUI3)* with high-copy *LEU2* plasmids containing the appropriate *SUI3* alleles derived from pAV1726 by site-directed mutagenesis, as described above (Table 7.1.1).

### 3.3.2 Biochemical assays using yeast cell extracts

Assays of β-galactosidase activity in whole cell extracts (WCEs) were performed as described previously (344). For Western analysis, WCEs were prepared by trichloroacetic acid extraction as previously described (345) and immunoblot analysis was conducted as previously described (210) using antibodies against eIF1/Sui1 (346) and Hcr1 (346). Enhanced chemiluminiscence (Amersham) was used to visualize immune complexes, and signal intensities were quantified by densitometry using NIH ImageJ software.

### 3.3.3 Biochemical analysis in the reconstituted yeast system

WT eIF1 and eIF1 variants Q31E, F108D, F108R, Q31AF108A, and K60E were expressed in bacterial strain BL21(DE3) Codon Plus cells (Agilent Technologies) and purified using the IMPACT system (New England Biolabs) as described previously (347). $His_6$-tagged WT eIF2 and eIF2 variants were overexpressed in yeast and purified as described (347). 40S subunits were purified as described previously from strain YAS2488 (347). Model mRNAs with sequences 5′-GGAA$[UC]_7$UAUG$[CU]_{10}$C-3′ and 5′-GGAA$[UC]_7$UUUG$[CU]_{10}$C-3′ were purchased from Thermo Scientific. Yeast $tRNA_i^{Met}$ was synthesized from a hammerhead fusion template using T7 RNA polymerase, charged with [$^{35}$S]-methionine, and used to prepare radiolabeled eIF2·GDPNP·[$^{35}$S]-Met-$tRNA_i$ ternary complexes ([$^{35}$S]-TC), all as previously described (347). Charged, unlabeled yeast Met-$tRNA_i^{Met}$ was purchased from tRNA Probes, LLC. For eIF1 binding competition experiments, WT eIF1 protein was labeled at its C-terminus with Cys-Lys-ε-fluorescein dipeptide, using the Expressed Protein Ligation system, as previously described (131).

TC dissociation rate constants ($k_{off}$) were measured by monitoring the amount of [$^{35}$S]-TC that remains bound to 40S·eIF1·eIF1A·mRNA (43S·mRNA) complexes over time, in the presence of excess unlabeled TC (chase), using a native gel shift assay to separate 40S-bound from unbound [$^{35}$S]-TC. 43S·mRNA complexes were preassembled for 2h at 26°C in reactions containing 40S subunits (20 nM), eIF1A (1 μM), eIF1 (WT or mutant variants, 1 μM), eIF5-G31R (1 μM) (where indicated), mRNA (10 μM), and [$^{35}$S]-TC (pre-assembled from 0.25 μM eIF2 (WT or mutant variants), 0.1 mM GDPNP, and 1 nM [$^{35}$S]-Met-$tRNA_i$) in 60 μL of reaction buffer (30 mM Hepes-KOH (pH 7.4), 100 mM potassium acetate (pH 7.4), 3 mM magnesium acetate, and 2 mM dithiothreitol). To initiate

each dissociation reaction, a 6 µL-aliquot of the preassembled 43S·mRNA complexes was mixed with 3 µL of 3-fold concentrated unlabeled TC chase (comprised of 2 µM WT eIF2, 0.3 mM GDPNP, and 0.9 µM Met-tRNA$_i$), representing a 300-fold excess over labeled TC in the final dissociation reaction, and incubated for the prescribed period of time. A converging time course was employed so that all dissociation reactions were terminated simultaneously by the addition of native-gel dye and loaded directly on a running native gel. The fraction of [$^{35}$S]-Met-tRNA$_i$ remaining in 43S complexes at each time point was determined by quantifying the 40S-bound and unbound signals by Phosphor Imaging, normalized to the ratio observed at the earliest time-point; and the data were fit with a single exponential equation (193).

TC association rates were measured by mixing [$^{35}$S]-TC with 40S·eIF1·eIF1A·mRNA complexes and quenching the binding reaction at various times by adding a 300-fold excess of unlabeled WT TC. Reactions were assembled as described above using 6 µL of sample and 3 µL of chase, and completed reactions were mixed with 2 µL of native gel dye before resolving 10 µL by native gel electrophoresis. As above, samples were loaded within minutes of one another on a running native gel. The $k_{obs}$ values were calculated by plotting the fraction of [$^{35}$S]-Met-tRNA$_i$ bound to 40S·eIF1·eIF1A·mRNA complexes against time and fitting the data with a single exponential equation. The resulting $k_{obs}$ values were plotted against the 40S subunit concentrations used in different experiments and the data were fit to a straight line. The slopes of these lines correspond to the second-order rate constants ($k_{on}$) for TC binding.

Fluorescence anisotropy measurements of equilibrium binding constants ($K_d$) for eIF1 binding to 40S·eIF1A complexes were performed using a T-format Spex Fluorolog-3 (J. Y. Horiba) as described previously (131). The excitation and emission wavelengths were 497 and 520 nm, respectively. The data were fit with a quadratic equation describing the competitive binding of two ligands to a receptor, as previously described (342).

## 3.4 RESULTS

### 3.4.1 Substitutions at the eIF2β:eIF1 interface decrease discrimination against suboptimal codons *in vivo*

Comparing the cryo-EM structures of the py48S-open and py48S-closed complexes reveals contacts between eIF2β and eIF1 restricted to the open complex (111) (Fig. 3.10.2 A) involving a potential stacking interaction of the side chains of Phe residues eIF2β-F217 and eIF1-F108. eIF2β-E198 and -Q221 appear to facilitate this interaction by stabilizing the position of eIF1-F108 via interactions with the protein backbone, and eIF2β-K216 and -G200 interact with eIF1-Q31. In an effort to disrupt the eIF2β-eIF1 interaction from each side of the interface, we introduced substitutions into eIF2β residues E198 and F217, and into eIF1 residues F108 and Q31. We predicted that perturbing the eIF2β:eIF1 interaction at each of these positions would destabilize eIF1 binding to the open complex, thus favoring isomerization to the closed state and increasing initiation at UUG codons.

Mutations introducing eIF2β substitutions were generated in a *SUI3* allele under its native promoter on a single copy *LEU2* plasmid, and the resultant mutant alleles were used to replace wild type *SUI3* on a *URA3* plasmid (plasmid shuffling) by counter-selection on

medium containing 5-fluoroorotic acid (5-FOA). Similarly, eIF1 mutations were generated in a *SUI1* allele under its native promoter on a single-copy *LEU2* plasmid, and the mutant *sui1* alleles were used to replace WT *SUI1* by plasmid-shuffling. Among the eIF1 substitutions, only eIF1-Q31E conferred a Slg$^-$ phenotype, which was restricted to 37º C (Fig. 3.10.2 B, row 4 vs. 1). The eIF2β-E198A substitution was lethal, preventing growth on 5-FOA medium, whereas the double substitution eIF2β-F217A/Q221A produced a slow-growth (Slg$^-$) phenotype (Fig. 3.10.2 C, row 2 vs. 1) at 30ºC. Previously, we established that eIF2β-F217A/Q221A is expressed at WT levels and forms a complex with eIF2α and eIF2γ comparably to WT eIF2β (111).

To examine the effects of the eIF2β and eIF1 substitutions on start codon selection, we first measured expression of matched *HIS4-lacZ* reporters containing either an AUG or UUG start codon. The eIF2β-F217A, eIF1-Q31A, eIF1-Q31K and eIF1-F108A substitutions all produced an ~2-4 -fold increase of the UUG:AUG ratio, while the double mutant eIF2β-F217A/Q221A and single mutants eIF1-Q31E and eIF1-F108D produced approximately order-of-magnitude increases compared to the WT strains (Fig. 3.10.2 D). The increased UUG:AUG ratios observed for eIF1-F108A, eIF1-F108D, and eIF2β-F217A/Q221A are in accordance with our previous observations on these variants (111). These hypoaccuracy phenotypes of multiple substitutions at the eIF2β:eIF1 interface support the notion that physical contacts between these factors observed exclusively in py48S-open help to stabilize eIF1 binding to the open, scanning conformation of the PIC. The greater increases in the *HIS4-lacZ* UUG:AUG initiation ratio conferred by replacing Q31 and F108 with acidic residues E or D versus Ala (Fig. 3.10.2 D) might result from

electrostatic repulsion with acidic residue E198 of eIF2β beyond the loss of eIF1:eIF2β contacts conferred by the corresponding Ala replacements. The much weaker hypoaccuracy phenotype of the eIF1-Q31K substitution might result from a new salt-bridge that can be formed between this altered residue and eIF2β-E198 (Fig. 3.10.2 D).

### 3.4.2 Substitutions at the eIF2β:eIF1 interface reduce discrimination against AUG codons in suboptimal context

There is evidence that the presence of eIF1 on the PIC disfavors initiation at AUG codons present in poor Kozak context as well as at non-AUG codons. Thus, eIF1 discriminates against the suboptimal sequence context of its own start codon in *SUI1* mRNA to autoregulate its cellular abundance (Martin-Marcos, 2011; Ivanov, 2010). Moreover, Sui$^-$ eIF1 mutations that elevate UUG initiation also reduce discrimination against the poor context of the *SUI1* AUG codon and increase eIF1 abundance (212, 342). Consistent with these previous findings, the eIF1-Q31A, Q31E, F108A, F108D and eIF2β-F217A/Q221A substitutions, which elevate UUG initiation on *HIS4-lacZ* mRNA (Fig. 3.10.2 D), all increase eIF1 abundance to some extent (Fig. 3.10.3 A). Importantly, the acidic substitutions eIF1-Q31E and eIF1-F108D with the strongest hypoaccuracy phenotypes (Fig. 3.10.2 D) also confer the largest increases in eIF1 levels, whereas Q31K, showing the smallest increase in UUG initiation, confers no significant increase in eIF1 abundance (Fig. 3.10.3 A). Furthermore, all of these substitutions except Q31A and Q31K increase expression of a *SUI1-lacZ* fusion containing the native, poor context of the *SUI1* AUG codon ($_{-3}$CGU$_{-1}$), but not that of a modified *SUI1-lacZ* fusion containing optimum context ($_{-3}$AAA$_{-1}$) (Fig. 3.10.3 B). In WT cells, the latter *SUI1$_{opt}$-lacZ* fusion is expressed

at ~2-fold higher levels than the native *SUI1-lacZ* fusion. Accordingly, the *SUI1$_{opt}$-lacZ:SUI1-lacZ* expression ratio is significantly diminished by eIF1-Q31A, Q31E, F108A, F108D and eIF2β-F217A/Q221A (Fig. 3.10.3 B). It is noteworthy that the strongest reductions in the *SUI1$_{opt}$-lacZ:SUI-lacZ* expression ratio resulted from the eIF1-Q31E, eIF1-F108D, and eIF2β-F217A/Q221A substitutions (Fig. 3.10.3 B), which also conferred the greatest increases in UUG initiation among this group of mutants (Fig. 3.10.2 D).

To support the conclusion that these substitutions reduce discrimination against AUGs in poor context, we asked whether they increase recognition of the suboptimal AUG codon of an upstream ORF, and thereby decrease expression of the downstream ORF encoded on the same mRNA. To this end, we assayed expression of *GCN4-lacZ* reporters containing a single, modified version of upstream ORF1 elongated to overlap the *GCN4* main ORF (el.uORF1). With the WT (optimal) context $_{-3}$AAA$_{-1}$ for el.uORF1, virtually all scanning ribosomes recognize its AUG codon (uAUG-1) and, because reinitiation at the downstream *GCN4* ORF following el.uORF1 translation is nearly non-existent, *GCN4-lacZ* expression is extremely low (348). In WT cells, replacing only the optimal A with U at the -3 position of el.uORF1 increases leaky scanning of uAUG-1 to produce an ~7-fold increase in *GCN4-lacZ* translation. Introducing the poor context of $_{-3}$UUU$_{-1}$ further increases leaky scanning, for an ~33-fold increase in *GCN4-lacZ* expression, and eliminating uAUG-1 altogether increases *GCN4-lacZ* expression by >100-fold (Fig. 3.10.3 C, column 1 (WT), rows 1-4 & Fig. 7.2.2 A,C). From these results, the percentages of scanning ribosomes that either translate el.uORF1 or bypass (leaky-scan) uAUG-1 and translate *GCN4-lacZ* instead can be calculated, revealing that >99.5%, ~93%, and ~67%

of scanning ribosomes recognize uAUG-1 in optimum, weak, and poor context, respectively, in WT cells (Fig. 3.10.3 C, cols. 4 & 7; see legend, Fig. 7.2.2 B).

Subjecting the mutants to this analysis revealed that eIF1 substitution Q31E reduces leaky scanning of uAUG-1, as indicated by significantly reduced *GCN4-lacZ* expression for the two reporters containing el.uORF1 with weak or poor context, but not for the el.uORF1 reporter with optimum context nor the uORF-less reporter (Fig. 3.10.3 C, cf. cols. 1-2, rows 1-4). However, calculating the percentages of ribosomes that recognize uAUG-1 revealed that eIF1-Q31E confers a substantial increase in recognition of uAUG-1 only for the poor-context (UUU) reporter, from ~67% to ~90%, while producing only a small increase for the weak-context (UAA) reporter, from ~93% to ~98% (Fig. 3.10.3 C, cf. cols. 7-8, rows 2-3; Fig 7.2.2 B, Q31E vs. WT, Weak & Poor). Similar results were obtained for the eIF1-F108D substitution, increasing recognition of uAUG-1 in poor-context (UUU) from ~67% to ~91%; while producing only a slight increase for weak-context (UAA), from ~93% to ~95%, and none for optimal context (AAA) (Fig. 3.10.3 C, cf. cols. 7 & 9, rows 1-3; Fig 7.2.2 B). Concordant results were also obtained for the eIF1-Q31A and eIF1-F108A substitutions (Fig. 7.2.3 A-D). Together, the results show that these eIF1 substitutions reduce discrimination against poor context, both for the *SUI1* AUG codon and uAUG-1 of *GCN4* mRNA, to an extent that generally parallels their effects in relaxing discrimination against UUG start codons at *HIS4*. We conclude that the eIF2β:eIF1 substitutions reduce discrimination against AUG codons in poor context and UUG start codons to similar extents.

### 3.4.3 Substitutions at the eIF2β:eIF1 interface reduce the rate of TC loading *in vivo*

Mutations that destabilize the open conformation of the PIC typically confer dual $Sui^-$ and $Gcd^-$ phenotypes, such as the K60E substitution in eIF1 Loop 1 that reduces eIF1 affinity for the 40S subunit (342). Because eIF1 promotes TC loading on the open conformation of the PIC (127), the reduced 40S occupancy of eIF1-K60E slows TC binding and confers derepression of *GCN4* translation *in vivo* ($Gcd^-$ phenotype) as revealed using a *GCN4-lacZ* reporter (Fig. 3.10.4 A, cols. 1-2). The 40S binding defect also enables inappropriate dissociation of eIF1-K60E from the scanning PIC at non-AUG codons to elevate UUG initiation (342). In contrast, substitutions in eIF1 Loop 2 that remove steric or electrostatic clashes with Met-$tRNA_i$ that impede transition to the closed/$P_{IN}$ state, but do not affect eIF1:40S association, increase UUG initiation without reducing the rate of TC loading *in vivo* ($Gcd^+$) (200).

None of the single eIF1 substitutions at the eIF2β:eIF1 interface significantly derepresses *GCN4-lacZ* expression (Fig. 3.10.4 A), and eIF2β-F217A increases expression only slightly (Fig. 3.10.4 B). However, substantial derepression of *GCN4-lacZ* occurred when eIF2β-F217A was combined with Q221A in the eIF2β-F217A/Q221A double substitution (Fig. 3.10.4 B), as reported previously (111). Derepression of *GCN4-lacZ* expression also was conferred when the eIF1 substitutions F108A and Q31A were combined (Fig. 3.10.4 C). The eIF1-F108A/Q31A double substitution also conferred a larger increase in the UUG:AUG initiation ratio compared to the two single substitutions (Fig. 3.10.4 D), as shown above for eIF2β-F217A/Q221A versus eIF2β-F217A alone (Fig. 3.10.2 D). Moreover, eIF1-Q31A/F108A conferred a greater increase in eIF1 abundance

compared to the corresponding single substitutions (Fig. 3.10.4 E), and it elevated expression of the native *SUI1-lacZ* fusion and reduced the *SUI1$_{opt}$-lacZ:SUI1-lacZ* expression ratio (Fig. 3.10.4 F) considerably more than did the eIF1-Q31A or eIF1-F108A single substitutions (Fig. 3.10.3 B). Together, these results indicate that combining the F108A and Q31A substitutions in eIF1 produces a synthetic reduction in discrimination against UUG codons and poor-context AUG codons in parallel with a derepression of *GCN4-lacZ* expression not observed in the corresponding single substitution mutants. Importantly, the latter defect was reversed by co-overexpressing all components of TC (Met-tRNA$_i$, eIF2α, eIF2β, and eIF2γ) from a high-copy plasmid (hc TC) (Fig. 3.10.4 G) as expected if the Gcd$^-$ phenotype of the eIF1-Q31A/F108A mutant results from slower TC loading that can be overcome by mass action at elevated TC concentrations. These genetic findings support the idea that substantially impairing the eIF1:eIF2β interface with multiple substitutions in either protein reduces the rate of TC loading to the open PIC conformation in addition to facilitating rearrangement from the open to closed conformation at poor start codons.

Because the Gcd$^-$ phenotype is conferred by eIF1 substitutions that weaken its binding to 40S subunits, we purified recombinant forms of the eIF1 variants and measured their 40S binding affinity by determining their ability to compete with fluorescently labeled WT eIF1 (eIF1$_{FL}$) for binding to purified 40S·eIF1A complexes *in vitro*. Pre-assembled 40S·eIF1A·eIF1$_{FL}$ complexes were challenged with increasing concentrations of unlabeled mutant or WT eIF1 proteins, and the fraction of bound eIF1$_{FL}$ remaining at each competitor concentration was determined by monitoring changes in fluorescence anisotropy (Fig.

3.10.5 A-B). In accordance with previous results (200, 342), the helix α1 variant K60E competed poorly with WT eIF1$_{FL}$ for binding to 40S·eIF1A complexes, increasing the eIF1 dissociation constant ($K_d$) by ~10-fold. By contrast, the eIF1-Q31E, F108D, and eIF1-Q31A/F108A mutants competed with WT eIF1$_{FL}$ indistinguishably from WT unlabeled eIF1, indicating no significant change in their affinity for 40S·eIF1A complexes (Fig. 3.10.5 B-C). These findings indicate that the decreased rate of TC binding inferred from the Gcd$^-$ phenotype of the F108A/Q31A double substitution, and the relaxed discrimination against poor initiation sites observed for this variant and the eIF1-F108D and eIF1-Q31E variants with single acidic substitutions, do not involve a weaker interaction of eIF1 with the 40S subunit.

We interpret the findings described thus far to indicate that the eIF1:eIF2β interface, found exclusively in the open conformation of the PIC, serves as a barrier to achieving the closed state, and that single substitutions at this interface that increase UUG initiation facilitate transition from the open to closed states at poor start codons by weakening this interface/barrier, without materially destabilizing the open conformation. The eIF1:eIF2β interface also facilitates TC binding during assembly of the open complex, and the more substantial weakening of the interface by multiple eIF1 or eIF2β substitutions reduces the rate of TC loading to the open complex in addition to conferring an even greater frequency of inappropriate transition to the closed state at poor start codons during scanning.

### 3.4.4 Substitutions at the eIF2β:eIF1 interface promote Met-tRNA$_i$ accommodation in the P$_{IN}$ conformation of the closed 48S PIC *in vitro*

To provide biochemical evidence that substitutions at the eIF2β:eIF1 interface facilitate isomerization to the closed/P$_{IN}$ state, we measured their effects on the rate of TC dissociation from PICs reconstituted from purified components *in vitro*. Partial 43S·mRNA PICs were reconstituted by incubating TCs pre-assembled with eIF2 (purified from yeast containing WT or mutant eIF2β subunits), [$^{35}$S]-Met-tRNA$_i$ and nonhydrolyzable GTP analog (GDPNP), with 40S subunits, saturating concentrations of WT or mutant eIF1, WT eIF1A, and an uncapped, unstructured model mRNA containing either an AUG or a UUG start codon [mRNA(AUG) or mRNA(UUG)]. The pre-assembled 43S·mRNA PICs were incubated with a chase of excess unlabeled WT TC for increasing time-periods, and the fraction of [$^{35}$S]-labeled TC remaining bound to the PIC was quantified after resolving 40S-bound and unbound fractions via native gel electrophoresis, from which the rate of dissociation ($k_{off}$) was calculated (Fig. 3.10.5 D). Previous work has indicated that TC bound in the open/P$_{OUT}$ conformation is unstable during electrophoresis and cannot be visualized. Thus, the measured rate of TC dissociation in these assays largely reflects the proportion of complexes in the P$_{IN}$ state and the stability of that state with either an AUG or UUG in the P site (193, 203).

In agreement with previous findings (193, 203, 214, 314), we observed that TC dissociates more rapidly from the mRNA(UUG) versus mRNA(AUG) PICs in reactions containing all WT components (Fig. 3.10.5 E-F, WT, Fig. 7.2.4; see representative data in Fig. 7.2.5 B), reflecting the reduced formation and relative instability of the P$_{IN}$ state at

near-cognate UUG versus AUG codons. The endpoint of the dissociation reaction is also seen to vary with start codon identity. At AUG codons, TC has dissociated from ~75% of PICs during the time course of our assay while the remaining ~25% of complexes are completely stable over this period (Fig. 3.10.5 E, G, WT). Those complexes from which no dissociation occurs are thought to reflect a shift to a hyperstable state distinct from the $P_{IN}$ state (203). The lower endpoint for complexes assembled on mRNA(UUG), of only ~9% of starting complexes (Fig. 3.10.5 E, G), presumably reflects the lower stability of the closed/$P_{IN}$ complex at near-cognate UUG codons, from which fewer complexes may enter the hyperstable state.

Interestingly, the double substitution eIF2β-F217A/Q221A results in dramatic stabilization of PICs at both AUG and UUG start codons. The degree of stabilization was such that off-rates could not be determined, because there was very little dissociation from these complexes on the time-scale of our assay (Fig. 3.10.5 E, F). The endpoints for both AUG and UUG start codons reflect almost no dissociation over 24 hours as compared to complexes before the addition of chase. This is similar to the effect seen previously for double substitutions of Watson-Crick base pairs in the ASL of Met-$tRNA_i$ (maintaining W:C pairing), which confer $Sui^-$ phenotypes (203). These data are consistent with the $Sui^-$ phenotype of eIF2β-F217A/Q221A, and suggest that this double eIF2β substitution shifts the equilibrium in favor of the closed/$P_{IN}$ state, which can be accessed with an elevated frequency at UUG codons. The fact that TC does not dissociate from ~80% of complexes containing eIF2β-F217A/Q221A further indicates that most complexes are able to enter

the highly stable state even with a near-cognate UUG codon in the P site (Fig. 3.10.5 E, G).

Related results were obtained for the purified eIF1 Sui$^-$ variants F108D and Q31E, as these substitutions decreased the rate of TC dissociation ($k_{off}$) by ~2.3- to 5.0-fold from PICs assembled on mRNA(UUG) (Fig. 3.10.5 H-I; see representative data in Fig. 7.2.5 C). They also decreased the extent of TC dissociation from these PICs as indicated by ~1.5- to 2.7-fold increases in the fraction of TC remaining bound to the PIC at the reaction end-points (Fig. 3.10.5 H, J), thus indicating increased formation of the hyper-stable complex at UUG codons. Using mRNA(AUG), eIF1-Q31E decreased the $k_{off}$ by ~2-fold, considerably smaller than the 5-fold reduction observed for mRNA(UUG); while eIF1-F108D had no significant effect on the rate of dissociation at AUG codons (Fig. 3.10.5 H-I), and both substitutions decreased the extent of TC dissociation at AUG codons (Fig. 3.10.5 H, J). These findings support the conclusion that substitutions in eIF1 or eIF2β that perturb the eIF2β:eIF1 interface in the open complex reduce a barrier to isomerization to the closed state, particularly at near-cognate UUG codons, in accordance with the increased initiation at UUG codons such mutations produce *in vivo*.

### 3.4.5 Substitutions at the eIF2β:eIF1 interface reduce the rate of Met-tRNA$_i$ loading to the open complex *in vitro*

As noted above, the eIF1-Q31A/F108A and eIF2β-F217A/Q221A double substitutions appear to reduce the rate of TC loading to the open conformation of the PIC, as indicated by elevated *GCN4-lacZ* reporter expression (Fig. 3.10.4 B-C). To support this

interpretation, we measured the kinetics of TC binding to 40S·eIF1A·eIF1·mRNA complexes assembled with either WT eIF1 or eIF1-Q31A/F108A using the gel mobility shift assay described above. WT TC preassembled with [$^{35}$S]-Met-tRNA$_i$ was added to initiate the reactions, which were terminated at each time point with excess unlabeled TC (Fig. 3.10.6 A). The pseudo-first-order rate constant ($k_{obs}$) was measured at different 40S concentrations to obtain the second-order rate constant ($k_{on}$) (193). Compared to WT eIF1, eIF1-Q31A/F108A confers an ~2-fold decrease in $k_{on}$ for mRNA(AUG) (Fig. 3.10.6 B-C, Representative data in Fig. 7.2.6 B). Using radiolabeled TCs assembled with eIF2 containing the eIF2β-F217A/Q221A mutant subunit, nearly identical ~1.5-fold decreases in $k_{on}$ were measured for both mRNA(AUG) and mRNA(UUG) compared to WT TCs (Fig. 3.10.6 D-E). The relatively smaller reduction in $k_{on}$ for eIF2β-F217A/Q221A versus eIF1-Q31A/F108A (Fig. 3.10.6 C, E) is in keeping with the weaker Gcd$^-$ phenotype observed for the former mutant *in vivo* (Fig. 3.10.4 B-C). The similar reduction in $k_{on}$ conferred by eIF2β-F217A/Q221A for PICs assembled with mRNA(AUG) or mRNA(UUG) fits with the notion that this eIF2β variant reduces the initial step of TC binding to the open conformation of the PIC, which is relatively insensitive to the nature of the start codon (193). These findings support the conclusion that the eIF1-Q31A/F108A and eIF2β-F217A/Q221A substitutions destabilize the open complex, reducing the rate of TC loading to this state while increasing rearrangement to the closed state at UUG codons during scanning.

### 3.4.6 eIF1-F108R increases discrimination against near-cognate UUG codons *in vivo* and favors the open/$P_{OUT}$ conformation of the PIC *in vitro*

Having found that the acidic substitution eIF1-F108D shifts the system to the closed complex and increases UUG initiation, and noting the presence of the essential acidic residue eIF2β E198 at the eIF1:eIF2β interface (Fig. 3.10.2 A, *left*), we asked whether substituting eIF1 F108 with the basic residue arginine would produce the opposite effect on the system versus eIF1-F108D by strengthening the eIF1:eIF2β interface through a new salt-bridge between eIF1 R108 and native eIF2β E198. Indeed, eIF1-F108R was found to partially suppress the elevated UUG initiation conferred by the dominant Sui$^-$ alleles *SUI5* and *SUI3-2*, encoding the eIF5-G31R and eIF2β-S264Y variants, respectively (349), thus conferring an Ssu$^-$ phenotype (Fig. 3.10.7 A-B, WT vs. *F108R*).

The eIF1-F108R substitution does not detectably alter eIF1 binding to 40S·eIF1A complexes *in vitro* (Fig. 3.10.5 B-C), suggesting that it reduces UUG initiation by impeding transition from the open to closed conformation of the PIC during scanning. Supporting this interpretation, eIF1-F108R increases the rate of TC dissociation from PICs reconstituted *in vitro* in the presence of the *SUI5* eIF5 variant, eIF5-G31R (Fig. 3.10.7 C-D). Consistent with our previous results (214), in reactions with WT eIF1 and eIF5-G31R, TC dissociates very little over the time course of the experiment from complexes containing an AUG or UUG start codon, yielding rate constants of only 0.10 $h^{-1}$ or 0.16 $h^{-1}$, respectively (Fig. 3.10.7 C-D), reflecting stabilization of the closed/$P_{IN}$ state by this eIF5 variant. Importantly, the TC dissociation rates for the complexes assembled

with eIF1-F108R and eIF5-G31R were increased by ~2.5-fold for mRNA(AUG) and ~4-fold for mRNA(UUG) compared to the $k_{off}$ values observed with WT eIF1 and eIF5-G31R (Fig. 3.10.7 C-D). The extent of TC dissociation was also elevated somewhat by eIF1-F108R, as reflected in lower reaction end-points (Fig. 3.10.7 E). These findings provide biochemical evidence that eIF1-F108R destabilizes the closed/$P_{IN}$ state at both AUG and UUG start codons, with a relatively stronger effect on the near-cognate triplet, thereby overriding the opposing effect of the eIF5-G31R substitution in preferentially enhancing the stability of the UUG complex (214). These findings help to account for the decreased utilization of UUG codons conferred by the F108R substitution *in vivo*.

### 3.4.7 Substitutions in eIF2β that disrupt contacts with the Met-tRNA$_i$ anticodon stem-loop (ASL) favor the closed/$P_{IN}$ complex and perturb TC binding to the open complex

Comparing the cryo-EM structures of py48S-open and py48S-closed complexes revealed contacts between eIF2β and the ASL of Met-tRNA$_i$ that are unique to the open state (Fig. 7.2.1 A-B). These include eIF2β residues S202, K170 and K214, which lie in proximity to the modified anticodon loop nucleotide $t^6$A37 (Fig. 3.10.8 A). In the closed complex, eIF2β is more compact and forms an alternative interface with the acceptor stem and D-loop, but not the ASL of Met-tRNA$_i$ (Fig. 7.2.1 B). The structures also reveal that the open conformation has a widened P site in which many of the canonical contacts between the ASL and the 40S body are missing (111). We therefore hypothesized that in the absence of these latter contacts, eIF2β residues interacting with the ASL have a role in

stabilizing Met-tRNA$_i$ binding to the PIC in the open conformation, leading to the prediction that substituting these residues would increase initiation at poor initiation codons *in vivo*.

A single alanine substitution at K170 did not significantly increase the *HIS4-lacZ* UUG:AUG initiation ratio (data not shown), whereas the eIF2β-K214A substitution elevated the UUG:AUG ratio 2.5-fold. However, combining Ala substitutions of either K170 or K214 with S202A resulted in stronger phenotypes, with eIF2β-K170A/S202A increasing the UUG:AUG ratio 3.0-fold and eIF2β-S202A/K214A producing a 9.1-fold increase (Fig. 3.10.8B), in accordance with our earlier findings on the -S202A/K214A variant (111). These results indicate that the loss of a positive charge at eIF2β residue 214 relaxes discrimination against near cognate start codons, particularly when a hydroxyl group is absent at residue 202, whereas loss of the basic side-chain of K170 has a similar but weaker effect on the system. The eIF2β-S202A/K214A variant also slightly increases eIF1 abundance (Fig. 3.10.3 A, *right*), and elevates expression of the *SUI1-lacZ* reporter containing the native, suboptimal AUG context while decreasing the *SUI1$_{opt}$-lacZ:SUI1-lacZ* ratio (Fig. 3.10.8 C), suggesting a modest reduction in discrimination against AUG codons in poor sequence context in addition to increasing UUG initiation. These results support the prediction that perturbing the eIF2β:Met-tRNA$_i$ interface specific to the open complex would increase the probability of rearrangement to the closed state at non-AUG codons and AUGs in poor context. The eIF2β-S202A/K214A variant is both expressed and forms a complex with eIF2α and eIF2γ at WT levels (111).

We also examined whether these eIF2β substitutions impair TC loading using the *GCN4-lacZ* reporter described above. While no Gcd$^-$ phenotype was seen for the single substitutions eIF2β-K214A or eIF2β-K170A (data not shown), the double substitutions eIF2β-K170A/S202A and eIF2β-S202A/K214A derepressed *GCN4-lacZ* expression by 2.2- and 3.7-fold, respectively (Fig. 3.10.8 D). Because TC loads onto the open state of the PIC, this TC loading defect, combined with the Sui$^-$ phenotypes described above, provide *in vivo* evidence that contacts between eIF2β residues and the ASL contribute to TC binding specifically to the open complex.

This last conclusion was further supported by biochemical analysis. Measuring the kinetics of TC dissociation revealed that eIF2β-S202A/K214A stabilized TC binding at both AUG and UUG codons, such that no dissociation was seen over the time-scale of our experiments (Fig. 3.10.8 G-I). This defect was coupled with a ~2.6- to 3-fold reduction in the rate of TC binding to PICs ($k_{on}$) with AUG or UUG start codons (Fig. 3.10.8 E-F). These results signify a destabilization of the open conformation, with attendant reduction in the rate of TC loading and a shift toward the closed conformation and subsequent rearrangement to the hyper-stable complex, on perturbing the eIF2β:Met-tRNA$_i$ ASL interface unique to the open complex.

### 3.4.8 eIF2β substitutions at the interface with the Met-tRNA$_i$ D-loop favor rearrangement to the closed/P$_{IN}$ complex with minimal perturbation of TC binding to the open complex

In addition to its contacts with the ASL described above, eIF2β forms differential contacts with the D-loop of Met-tRNA$_i$ in the open and closed conformations of the PIC. Interestingly, overlaying the two cryo-EM structures revealed extensive steric and electrostatic clashes between eIF2β in py48S-open and Met-tRNA$_i$ in py48S-closed, particularly involving eIF2β residues E189 and Q193 (Fig. 3.10.9 A, Fig. 7.2.1 C). These predicted clashes might indicate that eIF2β functions analogously to eIF1 during scanning to impede transition of Met-tRNA$_i$ to the P$_{IN}$ state of the closed complex before an AUG codon enters the P site. If so, then alanine substitutions of these residues should alleviate the clash, allowing Met-tRNA$_i$ to move into the closed/P$_{IN}$ state inappropriately at non-AUG codons. Indeed, we found that eIF2β-E189A, eIF2β-Q193A, and the double substitution eIF2β-E189A/Q193A elevated the UUG:AUG ratio between ~3.8- and ~6.1-fold (Fig. 3.10.9 B). Interestingly, substituting E189 with positively charged Arg conferred a considerably larger increase in UUG initiation compared to E189A (Fig. 3.10.9 B), whereas the eIF2β-Q193R substitution was lethal. eIF2β-E189R conferred a slow-growth phenotype (Fig. 3.10.2 C). Increased UUG initiation in response to the Ala substitutions of these eIF2β residues supports the notion that relieving a clash with Met-tRNA$_i$ at these positions removes a barrier to P$_{IN}$. That UUG initiation was further elevated by introducing a positive charge at position 189 suggests that introducing electrostatic attraction with Met-tRNA$_i$ can stabilize the closed/P$_{IN}$ state at UUG codons. Previously, we made similar findings for Ala and Arg substitutions in Loop-2 of eIF1, which is also predicted to clash

with the Met-tRNA$_i$ D-loop as a means of restricting transition to the closed complex (200). eIF2β-E189R slightly increased eIF1 abundance (Fig. 3.10.3 A, *right*) and modestly reduced the *SUI1$_{opt}$-lacZ:SUI1-lacZ* expression ratio (Fig. 3.10.9 D), indicating that eIF2β-E189R has a relatively small effect in promoting transition to the closed/P$_{IN}$ state at AUGs in poor context compared to its effect at UUG codons, suggesting that a codon:anticodon mismatch in the P-site is required for its effect on the system.

Of the eIF2β substitutions in this region, only Q193A produced a significant but modest 1.6-fold derepression of *GCN4-lacZ* expression. It is noteworthy that no derepression was observed for eIF2β-E189A/Q193A nor –E189R (Fig. 3.10.9 C) that conferred equal or greater increases in UUG initiation compared to eIF2β-Q193A (Fig. 3.10.9 B). The general lack of Gcd$^-$ phenotypes for substitutions at this interface suggests that the rate of TC loading is not impaired, supporting our hypothesis that these substitutions remove a clash with Met-tRNA$_i$ that impedes transition to the closed state during scanning without affecting the stability of the open state of the PIC.

Our interpretation of the *in vivo* phenotypes of the E189R variant is supported by biochemical analysis. Similar to the eIF1-F108D mutation, we observed that eIF2β-E189R lowered the TC off-rate ($k_{off}$) for UUG complexes by 2-fold while dramatically increasing the number of complexes from which Met-tRNA$_i$ does not dissociate (~5-fold) (Fig. 3.10.9 E-G). These results indicate a shift toward the closed conformation and subsequent rearrangement to the hyperstable state. The TC on-rates ($k_{on}$) for mRNA(AUG) and mRNA(UUG) were unaffected by the E189R substitution (Fig. 3.10.9 H-I), indicating that

this mutation does not confer a TC loading defect. This is consistent with the lack of Gcd$^-$ phenotype seen *in vivo* for this variant (Fig. 3.10.9 C) and supports our interpretation that the E189R substitution does not affect the stability of the open complex to which TC binds but only facilitates transition to the closed/$P_{IN}$ state.

## 3.5 DISCUSSION

In this report, we provide a combination of genetic and biochemical evidence that eIF2β plays a direct role in stimulating TC loading in the assembly of 43S PICs, and that it supports eIF1 function both in promoting scanning of the mRNA for AUG codons and in suppressing initiation at non-AUG start codons. These functions of eIF2β are mediated by its separate interactions with eIF1 and the ASL of Met-tRNA$_i$ found specifically in the open conformation of the PIC. We also provide evidence that eIF2β cooperates with eIF1 in preventing stable binding of Met-tRNA$_i$ in the P site at near cognate start codons through physical clashes predicted between both eIF1 and eIF2β in their locations found in the open complex and Met-tRNA$_i$ in the closed state. Hence, as demonstrated previously for eIF1 (12), eIF2β has a dual role in promoting the open scanning conformation of the PIC while impeding rearrangement to the closed conformation, which helps to restrict the transition from open to closed conformations of the PIC to AUG codons in optimum context.

### 3.5.1 eIF2β:eIF1 and eIF2β:Met-tRNA$_i$ interactions restricted to the PIC open conformation reduce the rate of TC loading and increase initiation at poor initiation sites

Interactions of eIF2β with eIF1 and Met-tRNA$_i$ were first observed in the py48S-open complex, thought to represent the scanning complex with Met-tRNA$_i$ not tightly locked into the P site, and these contacts were absent in the py48S-closed complex, thought to represent the PIC following AUG recognition (111). In the downward movement of the 40S head towards the 40S body in the transition from py48S-open to py48S-closed, the TC moves in conjunction with the 40S head, and the HTH domain of eIF2β is displaced from its interface with eIF1. At the same time, the eIF2β-HTH domain moves along the Met-tRNA$_i$ and is no longer engaged with the ASL in the closed complex. Because the eIF2β:eIF1 and eIF2β:ASL interactions occur exclusively in the open complex (Fig. 7.2.1 D), we hypothesized that they help to prevent eIF1 dissociation from the scanning PIC, and that they also anchor eIF1 in the position needed for its predicted clash with Met-tRNA$_i$ in the closed conformation to help impede rearrangement to the P$_{IN}$ state at non-AUG codons (111, 200).

We previously provided genetic support for this hypothesis (111), and have confirmed and extended the evidence here that substitutions in either eIF1 or eIF2β that perturb their interface in the open complex increase aberrant transition to the closed conformation at near-cognate UUG codons *in vivo*, elevating the UUG:AUG initiation ratio for a *HIS4-lacZ* reporter *in vivo*. We obtained biochemical evidence indicating that such substitutions reduce rates of TC dissociation from py48S PICs reconstituted at UUG

codons *in vitro*, indicating a shift to the closed conformation (and subsequent transition to a hyperstable state) that exhibits stable Met-tRNA$_i$ binding in the P site. Additional strong support for our hypothesis came from the finding that the eIF1-F108R substitution, designed to strengthen the eIF1:eIF2β interface, affected the system in ways opposite to the substitutions expected to weaken the interface, decreasing UUG initiation in a mutant rendered hypoaccurate by the eIF5-G31R variant (*SUI5*) and increasing the TC off-rate at UUG codons from py48S PICs assembled in the presence of eIF5-G31R.

An additional *in vivo* manifestation of eIF1 substitutions that favor the open-to-closed transition is the reduced leaky scanning of AUG start codons in poor context (212). We observed this phenotype here for both eIF1 and eIF2β substitutions that perturb their interface in py48S-open, as indicated by increased initiation at the eIF1 AUG present in the native, poor context, but not when context was optimized. In addition, they increased initiation at the AUG for the *GCN4* uORF1 when it resides in poor context, with attendant reduced initiation at the downstream *GCN4-lacZ* coding sequences. Thus, we have demonstrated that the eIF2β:eIF1 interface in the open complex is required for discrimination against poor context at AUG codons in addition to suppressing initiation at near-cognate triplets.

Substitutions at the eIF1:eIF2β interface also impair assembly of the PIC by reducing the rate of TC loading. This was revealed previously for the eIF2β-F217A/Q221A double mutant by its derepression of a *GCN4-lacZ* reporter (111), and was also observed here for the eIF1-F108A/Q31A double mutant. Translation of *GCN4* mRNA is regulated

by a delayed reinitiation mechanism wherein a decreased rate of TC loading blocks reinitiation at the uORFs and allows it to occur downstream at the *GCN4-lacZ* coding sequences instead (295). Here, we provided biochemical support for this interpretation by demonstrating a reduced rate of TC binding to py43S·mRNA complexes reconstituted *in vitro* using the eIF1-F108A/Q31A and eIF2β-F217A/Q221A variants. For the eIF1 mutant, we further established that the substitutions do not decrease eIF1 affinity for reconstituted 40S·eIF1A complexes *in vitro*, supporting a specific destabilization of the eIF1:eIF2β interface as the defect underlying both the assembly and accuracy phenotypes of these substitutions.

The fact that eIF1 and eIF2β substitutions that destabilize the open conformation simultaneously reduce the on-rate and decrease the off-rate for TC in reconstituted py48S complexes is consistent with previous biochemical results indicating that eIF1 in the open complex accelerates TC binding, but that TC is more tightly bound to closed complexes lacking eIF1 (that would exist following eIF1 release at the start codon) (127). The dual biochemical effects of the eIF1 and eIF2β substitutions in decreasing the rate of TC binding to the open conformation and reducing TC dissociation from the closed complex are paralleled by their dual *in vivo* phenotypes of derepressing *GCN4-lacZ* expression and elevating the *HIS4-lacZ* UUG:AUG initiation ratio. Hence, our combined genetic and biochemical analyses provide strong evidence that the eIF1:eIF2β interface both stabilizes the open conformation and opposes transition to the closed state during scanning (Fig. 7.2.7). These data further identify a direct molecular mechanism for the enhancement of TC recruitment by eIF1. Whereas single substitutions in eIF1 or eIF2β conferred elevated

UUG initiation, multiple substitutions were required to derepress *GCN4-lacZ* expression, suggesting that only a moderate perturbation of the eIF1:eIF2β interface is sufficient to increase the probability that eIF1 dissociates from the PIC during scanning, or to diminish eIF1's predicted clash with Met-tRNA$_i$ at non-AUG codons. However, a larger perturbation conferred by multiple substitutions in eIF1 or eIF2β is required to impair eIF1's enhancement of TC recruitment during PIC assembly.

Interactions between eIF2β and the ASL of Met-tRNA$_i$ were also revealed in the py48S-open complex that are absent in py48S-closed (111), suggesting that they too might specifically stabilize the open, scanning conformation of the PIC. Supporting this hypothesis, we demonstrated that the eIF2β-S202A/K214A double substitution, designed to perturb the eIF2β:ASL interface, confers the dual *in vivo* phenotypes of derepressed *GCN4-lacZ* expression and elevated UUG initiation, and produced the corresponding biochemical defects of a reduced TC on-rate and decreased TC off-rate, respectively, in reconstituted py48S PICs. Thus, the eIF2β:ASL interactions, like eIF1:eIF2β contacts, appear to enhance TC loading and scanning in the open conformation of the PIC, and disfavor rearrangement to the closed state, at near cognate start codons (Fig. 7.2.7).

It is noteworthy that the eIF2β substitutions at the eIF2β:ASL interface have a smaller stimulatory effect on initiation at the poor-context AUG codon of eIF1 compared to initiation at the UUG start codon of the *HIS4-lacZ* reporter, whereas the eIF1 substitutions at the eIF1:eIF2β interface, and other eIF1 substitutions at the 40S interface described previously (342), confer more comparable increases in initiation at poor-context

AUG and UUG codons. The enhanced recognition of poor-context AUGs produced by eIF1 substitutions could result from a decreased rate of scanning that increases the dwell-time at each triplet and provides more time to complete downstream reactions required for start codon selection ("slow-scanning" mechanism). Alternatively, it could result from an increased probability of rearrangement to the closed conformation at each triplet that elevates the likelihood of initiation at AUGs in poor context ("shifted equilibrium" mechanism). We have invoked the latter mechanism to account for the increased UUG initiation conferred by eIF1 substitutions that shift the equilibrium from the open to closed conformation of reconstituted PICs at UUG codons (Figs. 3.10.1 B & 7.2.6).

While our data do not directly address the mechanism by which the recognition of suboptimal start codons is enhanced, one way to account for our findings is to propose that both the eIF1 substitutions and eIF2β substitutions perturbing the eIF1:eIF2β interface increase UUG initiation by the "shifted equilibrium" mechanism, whereas they enhance initiation at poor-context AUG codons by the "slow scanning" mechanism mentioned above. The eIF2β substitutions at the interface with the Met-tRNA$_i$ ASL would also enhance UUG initiation by the "shifted equilibrium" mechanism, as they confer reduced TC off-rates in reconstituted PICs. However, because they do not reduce the occupancy of eIF1 in the scanning PIC, they cannot increase initiation at poor-context AUGs by the slow-scanning mechanism and can do so only by shifting the equilibrium between open and closed conformations of the PIC.

### 3.5.2 The predicted clash of eIF2β in the open conformation with Met-tRNA$_i$ in the closed state impedes initiation at near-cognate start codons

Previous studies have shown that eIF1 physically restricts full accommodation of Met-tRNA$_i$ in the P-site of the scanning complex in a manner that restricts rearrangement to the closed state until a perfect AUG:anticodon duplex is formed in the P site (200). The eIF1:Met-tRNA$_i$ clash was revealed by superimposing eIF1 in py48S-open with Met-tRNA$_i$ in py48S-closed. Similar modeling here revealed that the HTH domain of eIF2β in py48S-open clashes with Met-tRNA$_i$ in py48S-closed (Fig. 7.2.1 C), and we hypothesized that eIF2β cooperates with eIF1 in opposing Met-tRNA$_i$ binding in the P$_{IN}$ state prior to AUG recognition. Supporting this idea, Ala substitutions in eIF2β residues E189 and Q193 designed to diminish the predicted clash with the D-loop of Met-tRNA$_i$, increased UUG initiation *in vivo*. Interestingly, the E189R substitution conferred a stronger phenotype compared to eIF2β-E189A, suggesting that replacing electrostatic repulsion with attraction allows eIF2β to stabilize rather than impede Met-tRNA$_i$ binding in the closed complex and thereby overcome the destabilizing effect of a UUG:anticodon mismatch more effectively than is achieved by merely eliminating the acidic side chain of eIF2β E189. Consistent with this interpretation, the eIF2β-E189R substitution shifted the equilibrium to the closed state *in vitro*, reducing the TC off-rate from UUG codons in reconstituted py48S complexes. Importantly, however, the E189R substitution had no effect on *GCN4-lacZ* expression *in vivo*, nor on the on-rate of TC in reconstituted PICs. These findings can be explained by proposing that diminishing the predicted electrostatic clash of the eIF2β HTH domain with Met-tRNA$_i$ (or replacing it with electrostatic attraction) removes a barrier to the P$_{IN}$ state of the closed complex that increases UUG initiation, but it does not destabilize

the open complex and, hence, does not slow TC loading to the open conformation of the PIC. We recently obtained similar results for substitutions in Loop-2 of eIF1 that were designed to diminish its predicted clash with Met-tRNA$_i$ (200). Hence, we propose that the eIF2β HTH cooperates with eIF1 Loop-2 in the scanning PIC to establish a physical barrier to isomerization of Met-tRNA$_i$ to the P$_{IN}$ state of the closed complex (Fig. 7.2.7). Both eIF1 and the eIF2β HTH in their respective positions in the open complex are predicted to clash with different surfaces of the Met-tRNA$_i$ in the closed complex, and their contacts with one another in the open complex should mutually reinforce their distinct inhibitory interactions with Met-tRNA$_i$ (Fig. 7.2.1 D).

The eIF2β-E189R substitution does not increase initiation at the poor-context AUG start codon of eIF1 mRNA and in this respect, resembles the eIF2β substitutions at the interface with Met-tRNA$_i$-ASL in the open complex in selectively increasing initiation at UUG codons. Because eIF2β-E189R is not expected to perturb eIF1 binding to the open complex, it should not reduce the rate of scanning and overcome poor AUG context by the "slow-scanning" mechanism proposed above. The same reasoning should apply to the eIF1 loop-2 mutations that diminish the predicted clash with Met-tRNA$_i$. However, these substitutions were found to increase initiation at poor-context AUGs in addition to UUG codons (200). As a possible explanation for this discrepancy, it could be proposed that eIF1 loop-2 substitutions additionally perturb an interaction of eIF1 with the eIF3c-NTD (230) required for a WT rate of scanning. More work will be required to determine whether decreased discrimination against poor context can be achieved simply by removing a

barrier to the $P_{IN}$ state, which is sufficient to increase UUG initiation, or whether a kinetic defect in scanning is also required.

In summary, through a combination of genetics and biochemistry, we have identified three distinct interactions involving the eIF2β HTH domain that increase the accuracy of initiation. These involve separate eIF2β contacts with eIF1 and the Met-$tRNA_i$ ASL, which occur specifically in the open conformation of the PIC, and the predicted clash of the eIF2β HTH domain in its position in the open complex with Met-$tRNA_i$ in its location in the closed state—all of which impede the open-to-closed transition of the PIC and help limit initiation to AUG codons. These functions of eIF2β, and the analogous regulatory roles of eIF1, join a growing list of molecular interactions visualized in the py48S-open or py48S-closed PICs (111) that regulate scanning and start codon recognition, including interaction of the eIF1A N-terminal tail with the codon:anticodon duplex (206), interactions of Rps5/uS7 with eIF2α (314, 316), and Rps3-mRNA interactions at the mRNA entry channel of the 40S subunit (119). Our results further establish a direct role for eIF1 in promoting TC binding to the open complex through its interactions with the eIF2β-HTH, and we show that this aspect of PIC assembly is also stimulated by eIF2β-HTH interactions with the Met-$tRNA_i$ ASL.

## 3.6 SUPPLEMENTARY DATA

Supplementary Data are available at NAR online and included as Appendix B.

## 3.7 ACKNOWLEDGEMENT

We thank Jinsheng Dong, Fan Zhang, Jagpreet Nanda, and Jon Lorsch for advice and assistance in using the yeast reconstituted system, and all members of our laboratory and those of Jon Lorsch, Tom Dever, and Nick Guydosh for invaluable suggestions.

## 3.8 FUNDING

This work was supported by the Intramural Research Program of the National Institutes of Health.

## 3.9 AUTHOR CONTRIBUTIONS

L.M. generated mutants of eIF2β and carried out *in vivo* and *in vitro* analysis of eIF2β variants, including purification of mutant proteins, and helped to write the manuscript. A.T. generated mutants of eIF1 and carried out *in vivo* and *in vitro* analyses of eIF1 variants, including purification of mutant proteins, and helped to write the manuscript. A.G.H provided advice and project direction, supervised the work, and helped to write the manuscript.

## 3.10 FIGURES AND TABLES

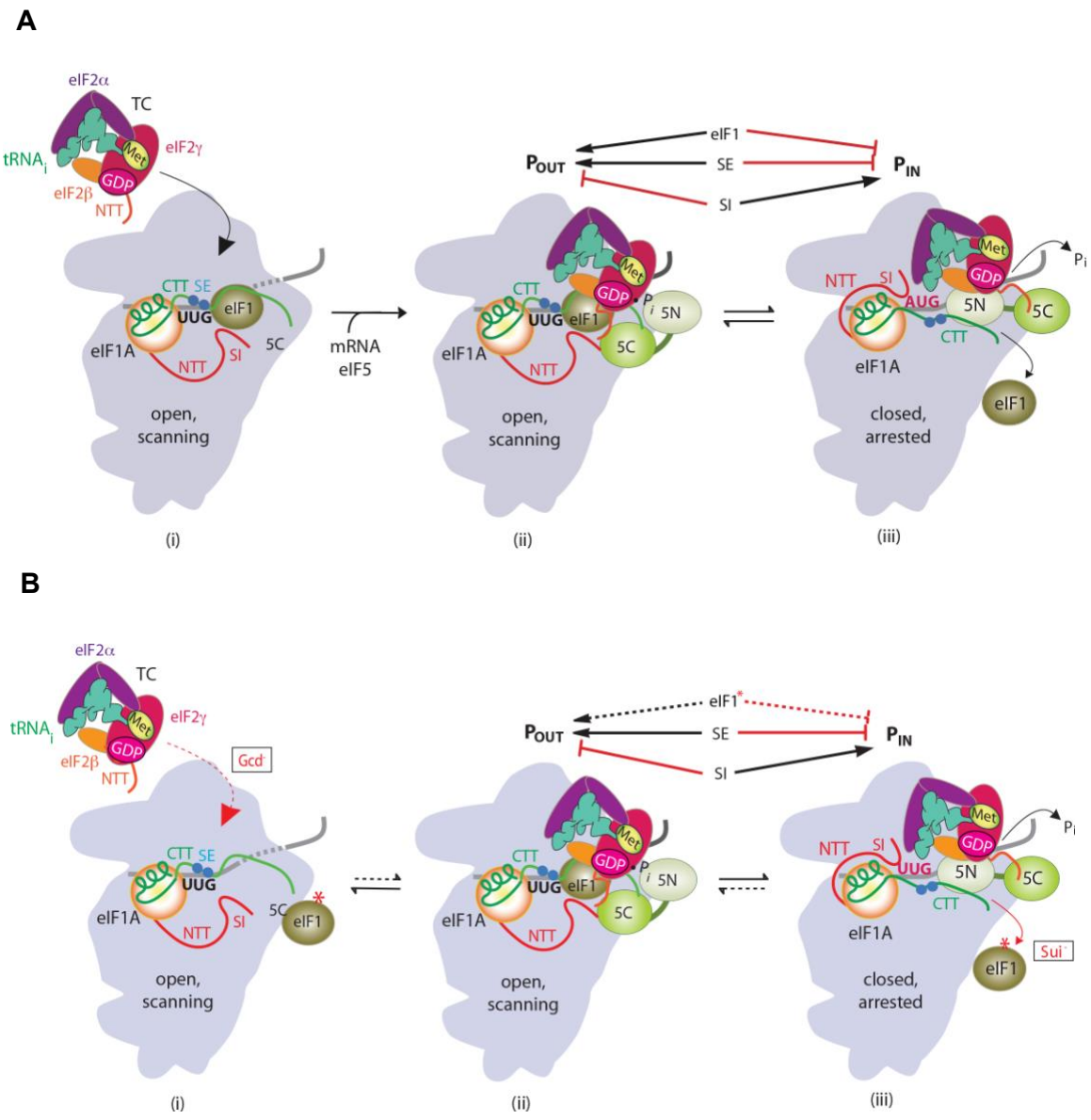

A
eIF2α
TC
tRNAi
eIF2γ
Met
GDP
eIF2β
NTT
CTT SE
UUG
eIF1
eIF1A
NTT
SI
5C
open,
scanning
(i)
mRNA
eIF5
eIF1
POUT
SE
SI
PIN
Met
GDP
Pi
5N
5C
(ii)
AUG
CTT
closed,
arrested
(iii)
B
Gcd-
eIF1*
Sui-

Fig. 1

**Figure 3.10.1. Model describing conformational rearrangements of the PIC during scanning and start codon recognition and the consequences of Sui$^-$ substitutions in eIF1**

**(A)** Assembly of the PIC, scanning and start codon selection in WT cells. (i) eIF1 and the scanning enhancer (SE) elements in the CTT of eIF1A stabilize an open conformation of the 40S subunit to which TC rapidly loads. (ii) The 43S PIC in the open conformation scans the mRNA for the start codon with Met-tRNA$_i$ bound in the P$_{OUT}$ state. The GAP domain in the N-terminal domain of eIF5 (5N) stimulates GTP hydrolysis by the TC to produce GDP-P$_i$, but release of P$_i$ is blocked. The unstructured NTT of eIF2β interacts with eIF1 to stabilize eIF1•40S association and the open conformation. (iii) On AUG recognition, the Met-tRNA$_i$ moves from the P$_{OUT}$ to P$_{IN}$ state, clashing with eIF1. Movement of eIF1 away from the P site disrupts its interaction with the eIF2β-NTT, and the latter interacts with the eIF5-CTD instead. eIF1 dissociates from the 40S subunit, and the eIF1A SE elements move away from the P site. The eIF5-NTD dissociates from eIF2 and interacts with the 40S subunit and the eIF1A CTT, facilitating P$_i$ release and blocking reassociation of eIF1 with the 40S subunit. (Above) The arrows summarize that eIF1 and the eIF1A SE elements promote P$_{OUT}$ and block the transition to the P$_{IN}$ state, whereas the scanning inhibitor (SI) elements in the NTT of eIF1A stabilizes the P$_{IN}$ state. (Adapted from Hinnebusch and Lorsch 2012, Nanda et al. 2013, Martin-Marcos et al. 2014.) **(B)** An eIF1 substitution that weakens that factor's binding to the 40S subunit destabilizes the open/P$_{OUT}$ conformation, reducing the rate of TC loading and increasing selection of near-cognate (UUG) start codons. (i) Aberrant dissociation of mutant eIF1 from the 40S subunit reduces the prevalence of the open/P$_{OUT}$ conformation, decreasing the rate of TC loading and

conferring the Gcd$^-$ phenotype (red dotted arrow). (ii, iii) Once TC eventually binds to the PIC and scanning commences, an increased frequency of mutant eIF1 dissociation from the open/$P_{OUT}$ conformation enables more frequent rearrangement to the $P_{IN}$ state at UUG codons, conferring the Sui$^-$ phenotype (red solid arrow).

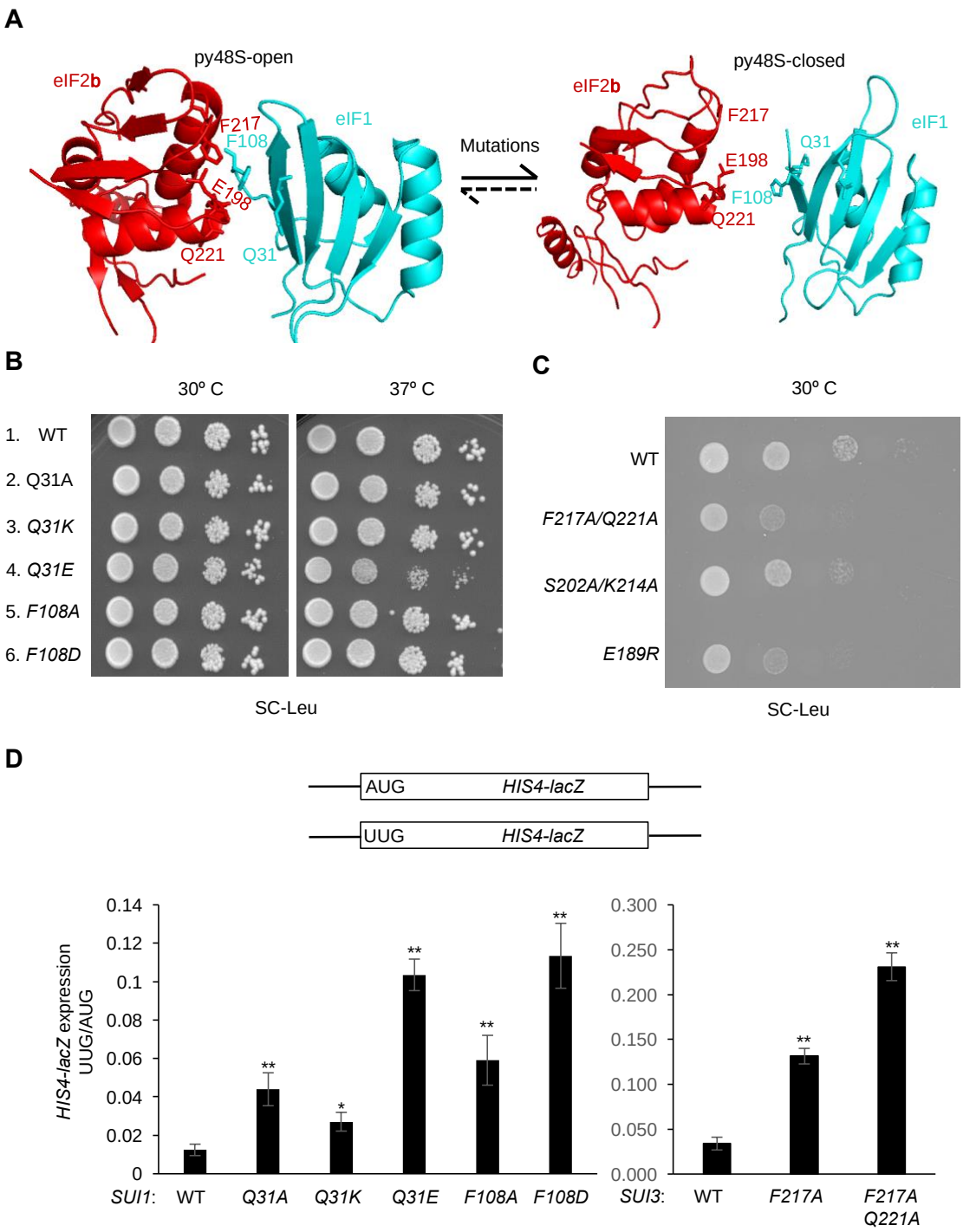
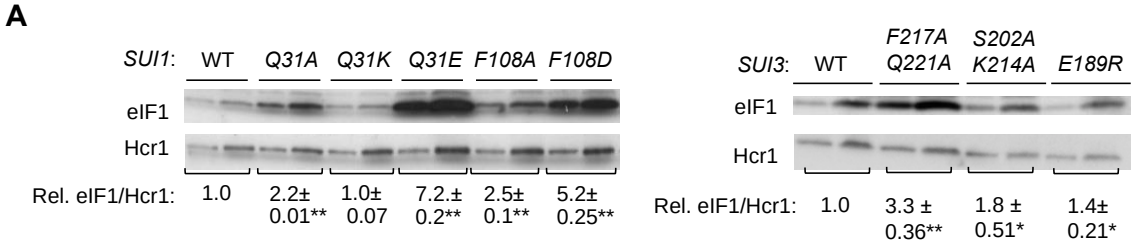

A
py48S-open
eIF2b
F217
F108
eIF1
E198
Q221
Q31
Mutations
py48S-closed
eIF2b
F217
Q31
eIF1
E198
F108
Q221
B
30° C
37° C
1. WT
2. Q31A
3. Q31K
4. Q31E
5. F108A
6. F108D
SC-Leu
C
30° C
WT
F217A/Q221A
S202A/K214A
E189R
SC-Leu
D
AUG HIS4-lacZ
UUG HIS4-lacZ
HIS4-lacZ expression UUG/AUG
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
SUI1: WT Q31A Q31K Q31E F108A F108D
0.300
0.250
0.200
0.150
0.100
0.050
0.000
SUI3: WT F217A F217A Q221A

Fig. 2

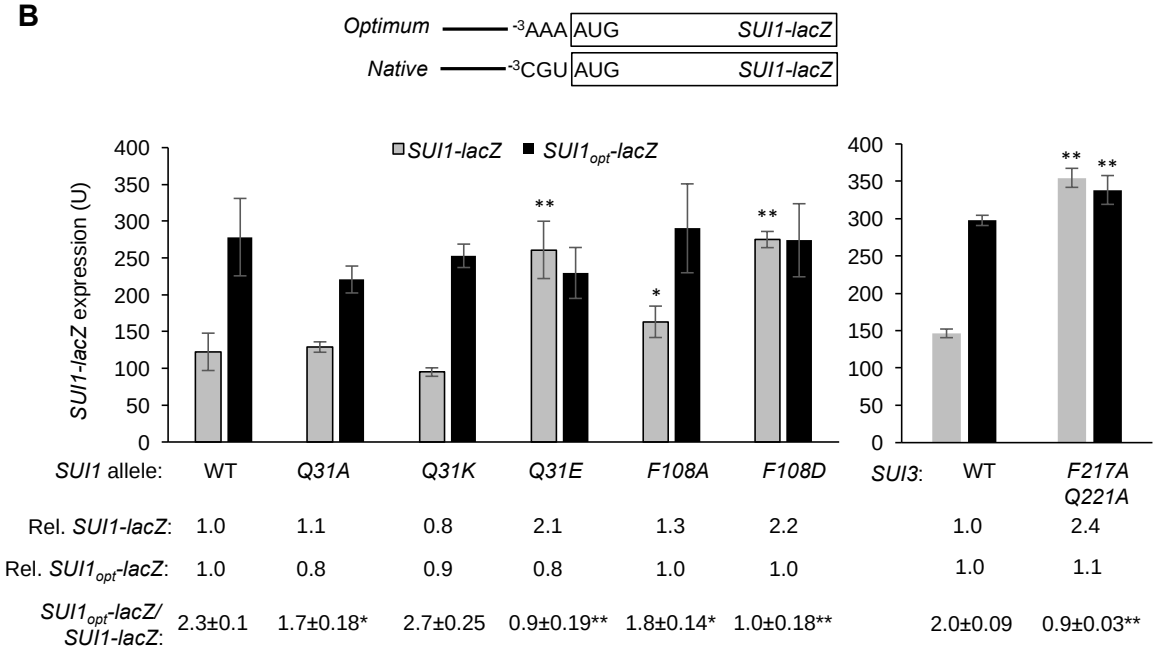
**Figure 3.10.2. Substitutions at the eIF2β:eIF1 interface of the open complex decrease discrimination against UUG initiation codons *in vivo***

**(A)** Interactions between eIF2β and eIF1 in the py48S-open and py48S-closed states of the PIC (111). eIF2β is shown in red, eIF1 in cyan. Residues substituted in this study are shown as sticks and labeled. Arrows indicate that substitutions disrupting these contacts favor the closed state by allowing eIF1 release. **(B)** Ten-fold serial dilutions of JCY03 derivatives with the indicated *SUI1* alleles were spotted on synthetic complete medium lacking leucine (SC-Leu) and incubated at 30˚C or 37˚C for 2-3d. **(C)** Ten-fold serial dilutions of KAY18 derivatives with the indicated *SUI3* alleles were spotted on SC-Leu and incubated at 30˚C for 2d. **(D)** *HIS4-lacZ* reporters with AUG or UUG start codons assayed to calculate UUG:AUG initiation ratios. Derivatives of *sui1Δ his4-301* strain JCY03 containing the indicated *SUI1* alleles or derivatives of KAY18 containing the indicated *SUI3* alleles, each harboring the appropriate *HIS4-lacZ* reporter, were cultured in synthetic dextrose minimal medium (SD) supplemented with His and Trp at 30˚C to $A_{600}$ of ~1.0, and β-galactosidase activities (in units of nanomoles of o-nitrophenyl-β-D-galactopyranoside cleaved per min per mg) were measured in whole cell extracts (WCEs). The ratio of expression of the UUG to AUG reporter was calculated from four to six different transformants, and the mean and SEMs were plotted. Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*$p < 0.05$; **$p < 0.01$).

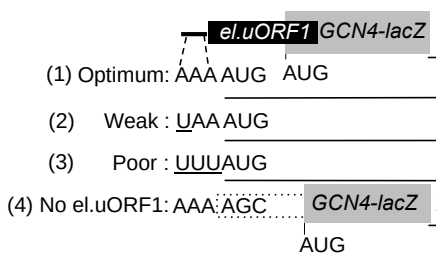
**A**




**B**




**C**




| | GCN4-lacZ expression(U) | | | % translating GCN4-lacZ | | | % translating el. uORF1 | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) WT | (2) Q31E | (3) F108D | (4) WT | (5) Q31E | (6) F108D | (7) WT | (8) Q31E | (9) F108D |
| (1) Optimum: AAA AUG AUG | 5±0.7 | 4±0.4 | 5±0.6 | 0.5 | 0.3 | 0.6 | 99.5 | 99.7 | 99.4 |
| (2) Weak : UAA AUG | 66±11 | 21±5** | 39±6** | 7 | 2 | 5 | 93 | 98 | 95 |
| (3) Poor : UUUAUG | 327±43 | 133±52** | 69±4** | 33 | 10 | 9 | 67 | 90 | 91 |
| (4) No el.uORF1: AAA AGC GCN4-lacZ AUG | 1000±150 | 1400±300 | 800±100 | >99 | >99 | >99 | <1 | <1 | <1 |

Fig. 3

**Figure 3.10.3. Substitutions at the eIF2β:eIF1 interface of the open complex decrease discrimination against AUG codons in suboptimal contexts *in vivo***

**(A)** Derivatives of *sui1Δ his4-301* strain JCY03 containing the indicated *SUI1* alleles or derivatives of KAY18 containing the indicated *SUI3* alleles were cultured in SD supplemented with His, Trp and Ura at 30˚C to $A_{600}$ of ~1.0, and WCEs were subjected to western blot analysis using antibodies against eIF1 and Hcr1 (loading control). Two amounts of each extract differing by a factor of two were loaded in successive lanes. eIF1 western signals were normalized to those for Hcr1 and mean values (±SEM) were calculated from seven biological replicates. **(B**) The same strains as in (A) but harboring sc plasmids (pPMB24 or pPMB25) with *SUI1-lacZ* fusions with native suboptimal ($_{-3}$CGU$_{-1}$) or optimum ($_{-3}$AAA$_{-1}$) AUG contexts were cultured and assayed for β-galactosidase activities as in Fig. 3.10.2 B. Mean expression levels and SEMs were calculated from three transformants, and relative (Rel.) mean expression levels normalized to that of the WT strain are listed below, along with expression ratios for the *SUI1-lacZ* versus *SUI1-opt-lacZ* reporters. **(C)** Transformants of JCY03 harboring WT *SUI1*, *sui1-Q31E* or *sui1-F108D* and el.uORF1 *GCN4-lacZ* reporters (pC3502, pC3503 or pC4466) containing the depicted optimum (row 1), weak (row 2) or poor (row 3) context of uAUG-1, or an uORF-less *GCN4-lacZ* reporter with a mutated uAUG-1 (pC3505, row 4), were assayed for β-galactosidase activities as in Fig. 3.10.2 B. Mean expression values with SEMs were determined from six transformants (columns 1, 2 and 3). The percentages of scanning ribosomes that translate el.uORF1 (columns 7, 8 and 9) or leaky-scan uAUG-1 and translate *GCN4-lacZ* instead (columns 4, 5 and 6) were calculated from results in columns 1, 2 and 3 by comparing the amount of expression observed for each uORF-containing

reporter to the uORF-less construct, as described in Fig. 7.2.2 B-C. (A-C) Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*$p < 0.05$; **$p < 0.01$).

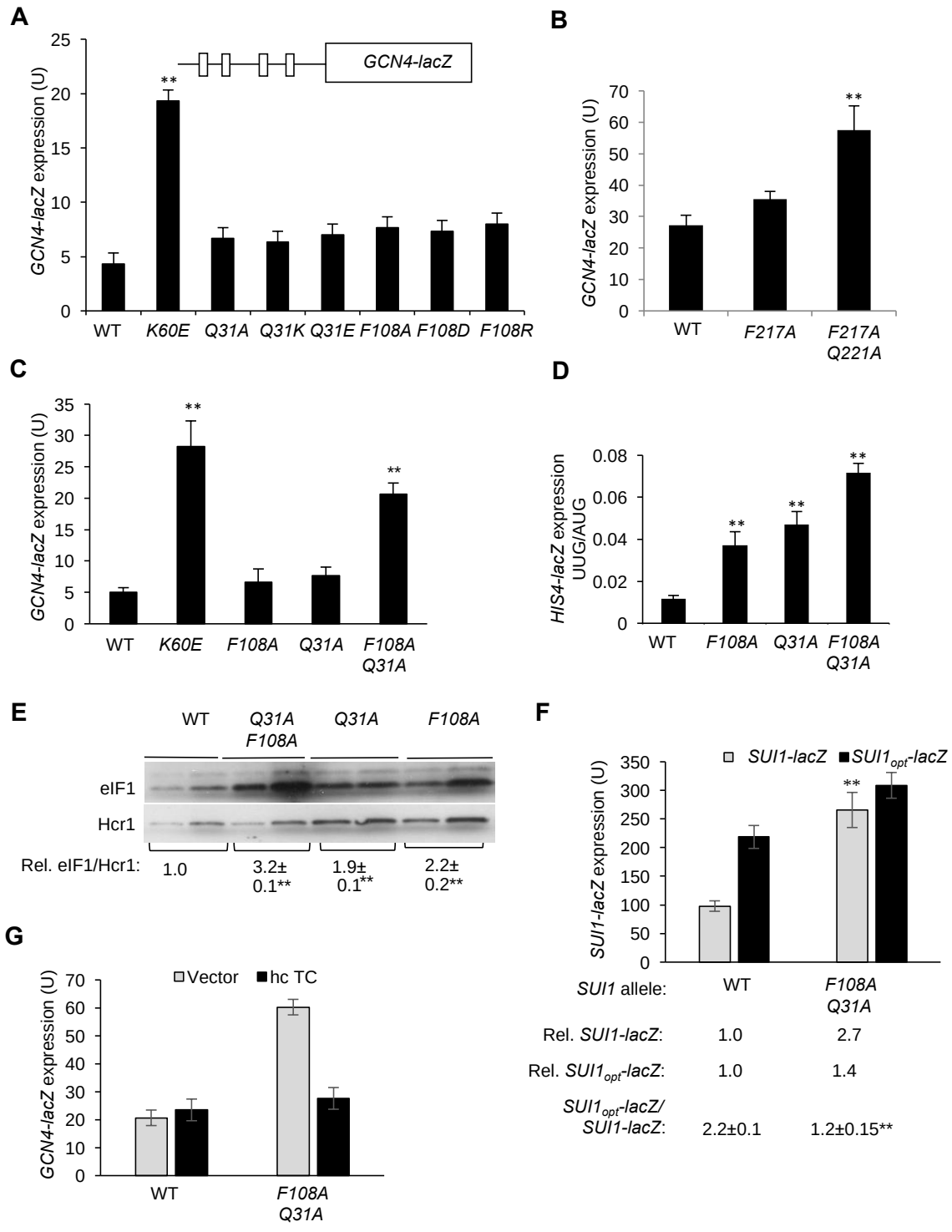
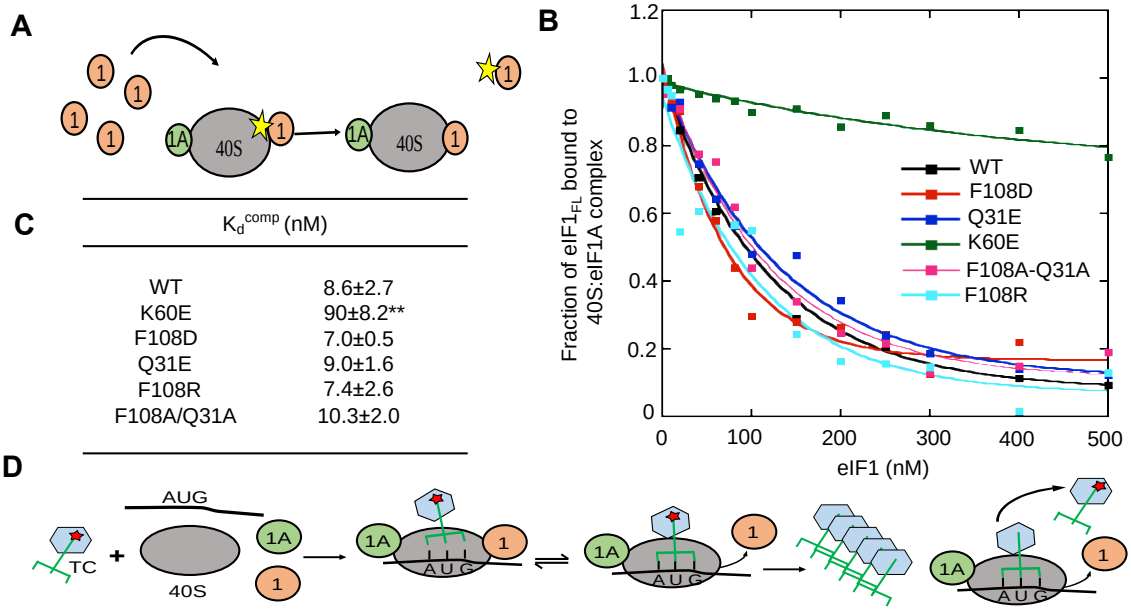

A
GCN4-lacZ
GCN4-lacZ expression (U)
WT
K60E
Q31A
Q31K
Q31E
F108A
F108D
F108R
B
GCN4-lacZ expression (U)
WT
F217A
F217A Q221A
C
GCN4-lacZ expression (U)
WT
K60E
F108A
Q31A
F108A Q31A
D
HIS4-lacZ expression UUG/AUG
WT
F108A
Q31A
F108A Q31A
E
WT
Q31A F108A
Q31A
F108A
eIF1
Hcr1
Rel. eIF1/Hcr1:
1.0
3.2± 0.1**
1.9± 0.1**
2.2± 0.2**
F
SUI1-lacZ
SUI1opt-lacZ
SUI1-lacZ expression (U)
SUI1 allele:
WT
F108A Q31A
Rel. SUI1-lacZ:
1.0
2.7
Rel. SUI1opt-lacZ:
1.0
1.4
SUI1opt-lacZ/ SUI1-lacZ:
2.2±0.1
1.2±0.15**
G
Vector
hc TC
GCN4-lacZ expression (U)
WT
F108A Q31A

Fig. 4

**Figure 3.10.4. Double substitutions at the eIF2β:eIF1 interface of the open complex derepress *GCN4-lacZ* expression and decrease discrimination against the poor context of the eIF1 AUG codon *in vivo***

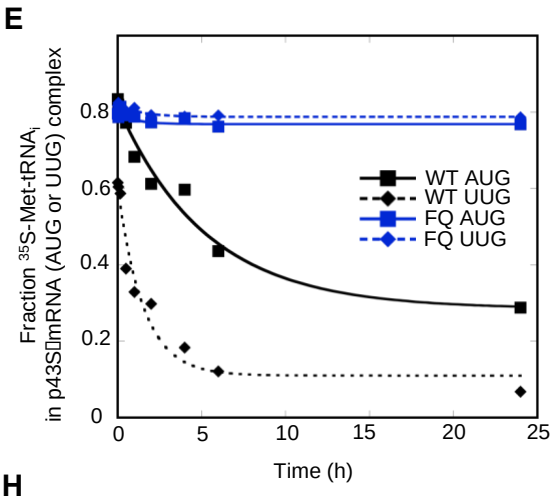
**(A)** Transformants of JCY03 containing the indicated *SUI1* alleles and the WT *GCN4-lacZ* reporter (depicted schematically) on plasmid p180 were assayed for β-galactosidase activities as in Fig. 3.10.2 B. Mean expression levels and SEMs calculated from six to eight transformants of each strain are plotted. **(B)** Derivatives of KAY18 containing the indicated *SUI3* alleles and the WT *GCN4-lacZ* reporter were assayed for β-galactosidase activities as in Fig. 3.10.2B. **(C)** Derivatives of JCY03 containing the indicated *SUI1* alleles and WT *GCN4-lacZ* reporter were assayed for β-galactosidase activities as in Fig. 3.10.2 B. **(D)** Transformants of JCY03 containing the indicated *SUI1* alleles and *HIS4-lacZ* reporters were assayed for β-galactosidase activities as in Fig. 3.10.2 B. **(E)** Derivatives of *sui1Δ his4-301* strain JCY03 containing the indicated *SUI1* alleles were subjected to western blot analysis as in Fig. 3.10.3 A. The eIF1 western signals were normalized to those for Hcr1 and mean values (±SEM) were calculated from three biological replicates. **(F**) Derivatives of *sui1Δ his4-301* strain JCY03 containing the indicated *SUI1* alleles and harboring plasmids (pPMB24 or -25) with *SUI1-lacZ* fusions with native suboptimal ($_{-3}$CGU$_{-1}$) or optimum ($_{-3}$AAA$_{-1}$) AUG contexts were cultured and assayed for β-galactosidase activities as in Fig. 3.10.2 B. Mean expression levels and SEMs were calculated from six transformants, and relative (Rel.) mean expression levels normalized to that of the WT strain are listed below, along with expression ratios for the *SUI1-lacZ* versus *SUI1-opt-lacZ* reporters. **(G)** Transformants of JCY03 containing the indicated *SUI1* alleles on a sc plasmid, either hc TC plasmid (p4835) or empty vector (YCplac112), and the WT *GCN4-*

*lacZ* reporter were assayed for β-galactosidase activities as in Fig. 3.10.2 B. (A-G) Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*$p < 0.05$; **$p < 0.01$).




C

| | $K_d^{comp}$ (nM) |
|---|---|
| WT | 8.6±2.7 |
| K60E | 90±8.2** |
| F108D | 7.0±0.5 |
| Q31E | 9.0±1.6 |
| F108R | 7.4±2.6 |
| F108A/Q31A | 10.3±2.0 |

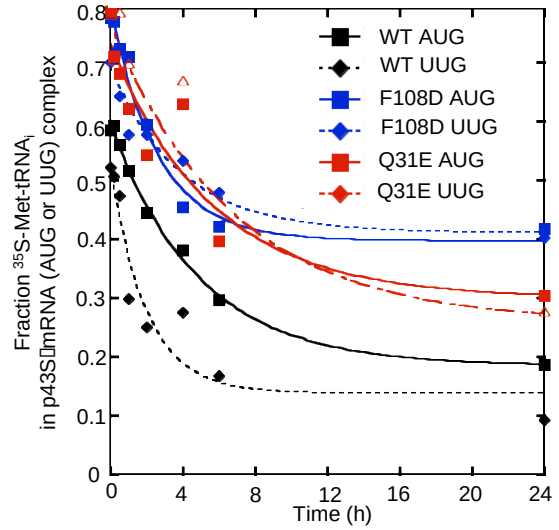



F

| | $k_{off}$ ($h^{-1}$) | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.66±0.07 | 0.17±0.01 |
| F217A/Q221A | ND | ND |

G

| | End-point | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.09±0.01 | 0.25±0.02 |
| F217A/Q221A | 0.77±0.01** | 0.77±0.02** |




I

| | $k_{off}$ ($h^{-1}$) | |
|---|---|---|
| | mRNA | |
| eIF1 | UUG | AUG |
| WT | 0.60±0.06 (5) | 0.19±0.005 (5) |
| *F108D* | 0.27±0.06** (3) | 0.18±0.1 (3) |
| *Q31E* | 0.12±0.03** (4) | 0.10±0.01**(4) |

J

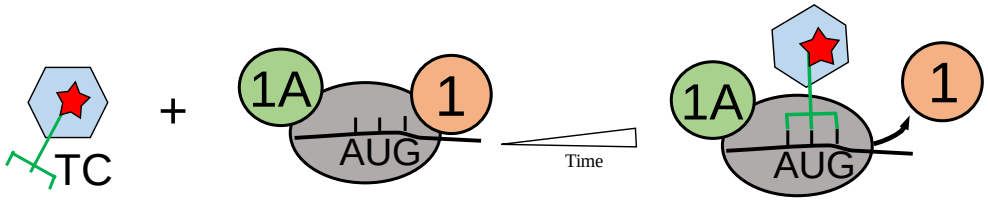
| | End-point | |
|---|---|---|
| | mRNA | |
| eIF1 | UUG | AUG |
| WT | 0.12±0.02 (5) | 0.18±0.002 (5) |
| *F108D* | 0.34±0.06** (3) | 0.36±0.08* (3) |
| *Q31E* | 0.24±0.18** (4) | 0.34±0.03** (4) |

Fig. 5

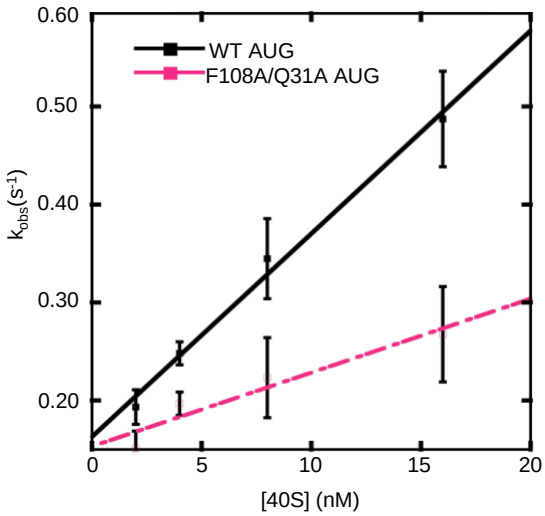
**Figure 3.10.5. Substitutions at the eIF2β:eIF1 interface of the open complex do not impair eIF1:40S interaction but enhance the closed/$P_{IN}$ conformation of py48S PICs at UUG codons *in vitro***

**(A-C)** Measurement of eIF1 binding constants. Fluorescein-labeled WT eIF1 (5 nM) was pre-bound to 40S subunits (15 nM) in the presence of eIF1A (1μM), mixed with increasing concentrations of unlabeled WT eIF1, eIF1-K60E, eIF1-F108D, eIF1-Q31E, eIF1-F108R or eIF1-F108D/Q31E, and the change in fluorescence anisotropy was measured (A). One of two replicate experiments is shown (B), from which mean $K_d^{comp}$ values and average deviations were calculated (C). **(D-J)** Measurement of TC dissociation kinetics. Partial 48S complexes were assembled with radiolabeled TC containing WT or mutant eIF2β, eIF1A, model mRNA containing an AUG or UUG start codon, and WT or mutant eIF1 proteins; chased with excess unlabeled TC for increasing periods of time; and the fraction of labeled Met-$tRNA_i$ bound to the PIC at each time-point was determined by EMSA (D). A representative plot of the fraction of [$^{35}$S]-Met-$tRNA_i$ incorporated into p43S·mRNA complexes is plotted as a function of time for eIF2β-F217A/Q221A (E) from which mean rate constants (F) and end points (G) (with SEMs) were calculated. Similar plots were created for eIF1 variants F108D and Q31E (H) and were used to calculate mean rate constants (I) and end points (with SEMs) (J) from between three and five replicate experiments (numbers in parentheses). **(F-J)** Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*, $p<0.05$; **, $p<0.01$).

A



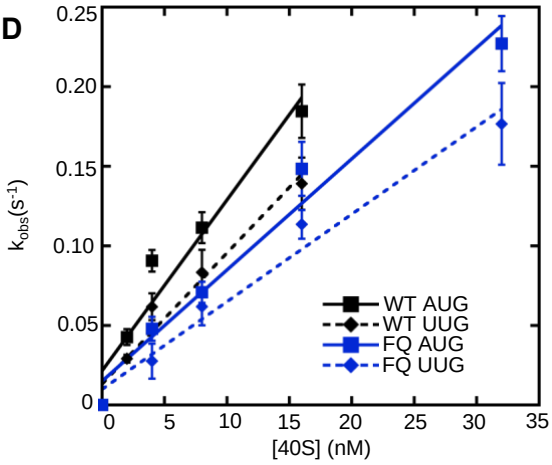

B




C

| | $k_{on}$ ($M^{-1}s^{-1}$) (×$10^6$) |
|---|---|
| eIF1 | |
| WT | 18.3±1.5 |
| F108A/Q31A | 8.9±1.3* |

D



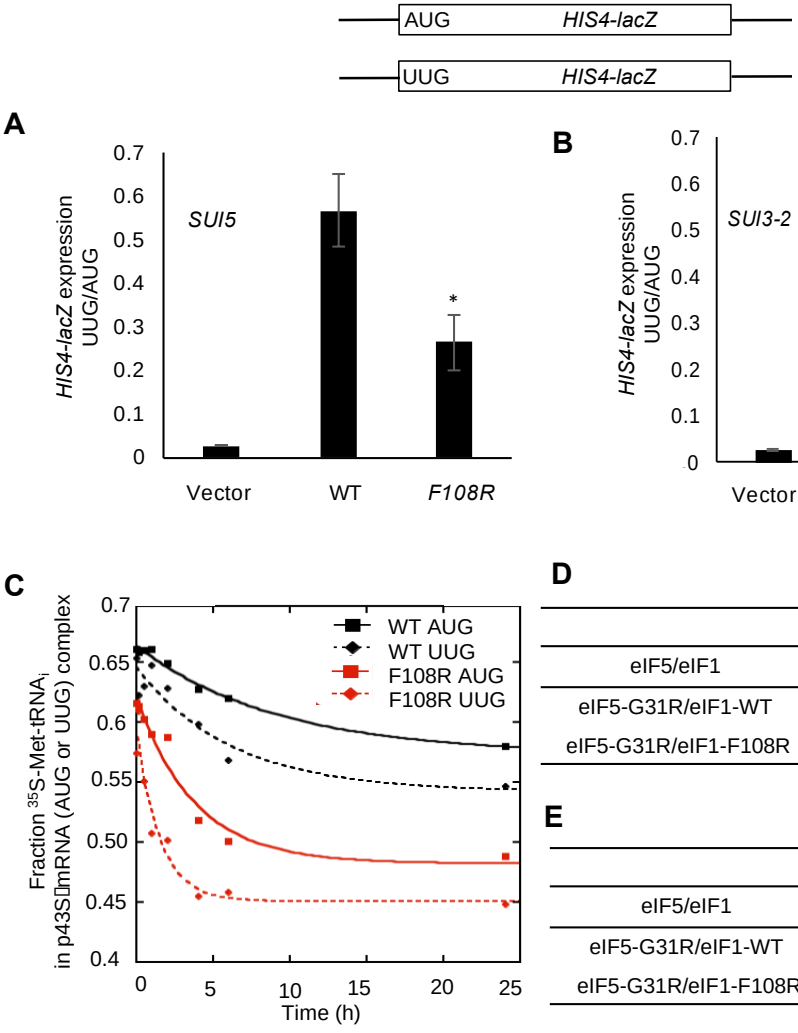

E

| | $k_{on}$ ($M^{-1}s^{-1}$) (×$10^6$) | |
|---|---|---|
| | mRNA | |
| eIF2β | AUG | UUG |
| WT | 10.73±1.0 | 8.23±1.2 |
| F217A/Q221A | 6.98±0.5* | 5.50±0.9* |

Fig. 6

**Figure 3.10.6. Double substitutions at the eIF2β:eIF1 interface of the open complex reduce rates of TC binding *in vitro***

**(A)** Schematic for measurement of TC association kinetics. WT TC preassembled with [$^{35}$S]-Met-tRNA$_i$ was mixed with pre-formed 40S·eIF1A·eIF1·mRNA complexes, incubated for increasing times, and reactions were terminated at each time-point with a chase of excess unlabeled TC. The fraction of labeled Met-tRNA$_i$ bound to the PIC at each time-point was determined by EMSA. The pseudo-first-order rate constant ($k_{obs}$) was measured at different 40S concentrations to obtain the second order rate constant ($k_{on}$). **(B-C)** Determination of TC $k_{on}$ values as described in (A) for WT or eIF1-F108A/Q31A mutant and mRNA(AUG). The mean $k_{obs}$ values were determined from at least three independent experiments at each 40S concentration. The resulting mean $k_{obs}$ values are plotted against each 40S concentration with error bars representing SEMs (B). $k_{obs}$ values from each of three independent experiments were plotted against 40S concentration in order to calculate $k_{on}$, and mean $k_{on}$ values (with SEMs) are shown (C). **(D-E)** Determination of TC $k_{on}$ values as described in (A-C) for WT eIF2 or eIF2 containing the eIF2β-F217A/Q221A (FQ) variant for partial 43S·mRNA complexes containing mRNA(AUG) or mRNA(UUG). The mean $k_{obs}$ values determined from three independent experiments at each 40S concentration were plotted versus the 40S concentration (D), and the mean calculated $k_{on}$ values determined from three independent experiments are shown (E). Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*), $p<0.05$.




**D**

| | $k_{off}$ ($h^{-1}$) | |
|---|---|---|
| | mRNA | |
| eIF5/eIF1 | UUG | AUG |
| eIF5-G31R/eIF1-WT | 0.16±0.009 | 0.10±0.03 |
| eIF5-G31R/eIF1-F108R | 0.66±0.06** | 0.26±0.02** |

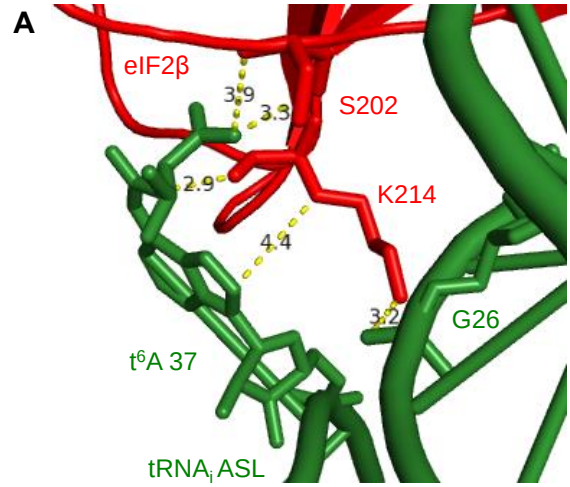
**E**

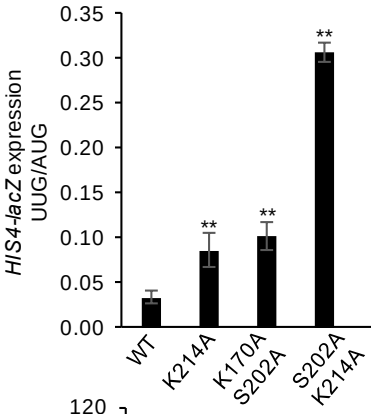
| | End-point | |
|---|---|---|
| | mRNA | |
| eIF5/eIF1 | UUG | AUG |
| eIF5-G31R/eIF1-WT | 0.54±0.01 | 0.55±0.01 |
| eIF5-G31R/eIF1-F108R | 0.47±0.01** | 0.50±0.01* |

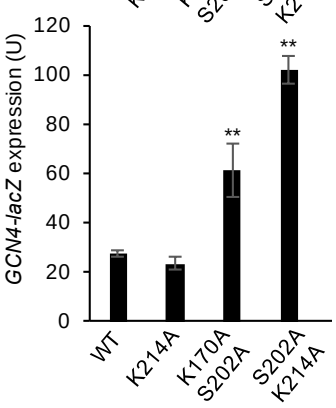
Fig. 7

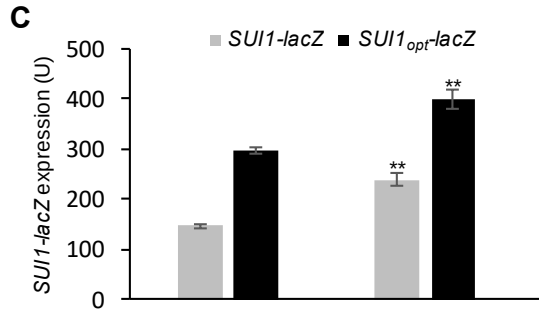
**Figure 3.10.7. eIF1 substitution F108R at the eIF2β:eIF1 interface of the open complex increases discrimination against UUG start codons *in vivo* and disfavors $P_{IN}$ at UUG start codons *in vitro***

**(A, B)** *HIS4-lacZ* reporters with AUG or UUG start codons were assayed as in Fig. 3.10.2 D (except that Trp was omitted from the medium) to calculate the UUG:AUG initiation ratio in the presence of the dominant $Sui^-$ allele *SUI5* on plasmid p4281 (A) or *SUI3-2* on plasmid p4280 (B). **(C-E)** TC dissociation kinetics were assayed as in Fig. 3.10.5 D-G for partial 43S·mRNA complexes containing mRNA(AUG) or mRNA(UUG) and assembled with WT eIF1 or eIF1-F108R in the presence of the eIF5-G31R variant (*SUI5*). Representative curves selected from three independent experiments are shown (C) from which mean rate constants (D) and end points (E) (with SEMs) were calculated. **(A-E)** Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (*, $p<0.05$; **, $p<0.01$).

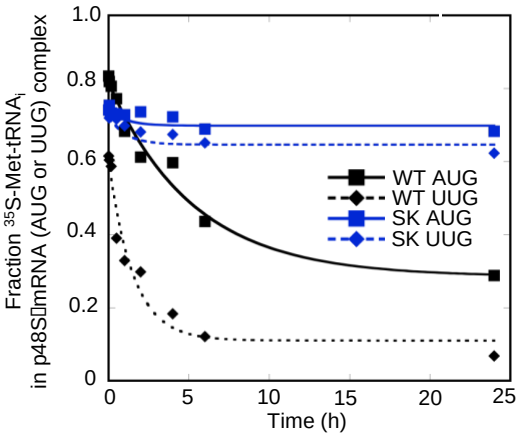


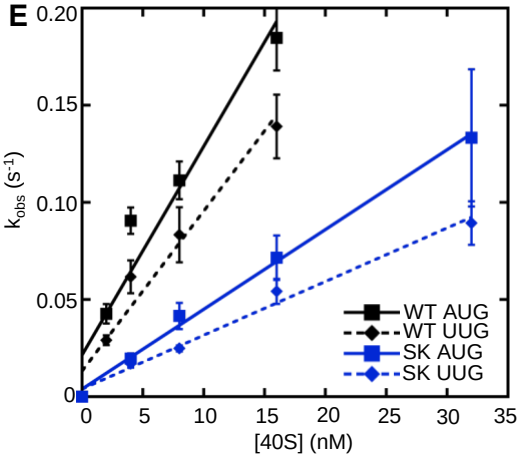




| SUI1 allele: | WT | S202A/K214A |
|---|---|---|
| Rel. SUI1-lacZ: | 1.0 | 1.6 |
| Rel. SUI1opt-lacZ: | 1.0 | 1.3 |
| SUI1opt-lacZ/ SUI1-lacZ: | 2.0±0.09 | 1.7±0.12 |




F

| | kon (s-1) | |
|---|---|---|
| | mRNA | |
| eIF2β | AUG | UUG |
| WT | 10.73±1.0 | 8.23±1.2 |
| S202A/K214A | 4.11±1.1** | 2.76±0.4** |

H

| | koff (h-1) | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.66±0.07 | 0.17±0.01 |
| S202A/K214A | ND | ND |

I

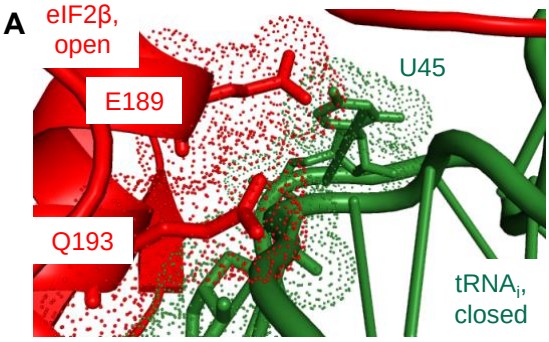
| | End-point | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.09±0.01 | 0.25±0.02 |
| S202A/K214A | 0.61±0.02** | 0.62±0.03** |

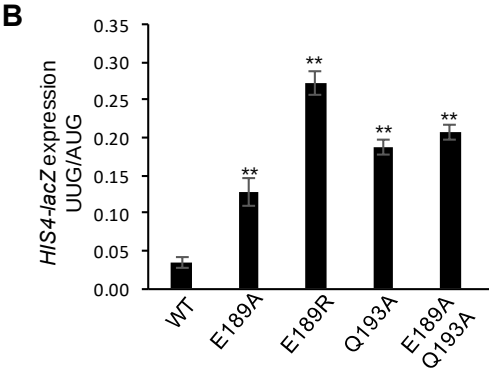
Fig. 8

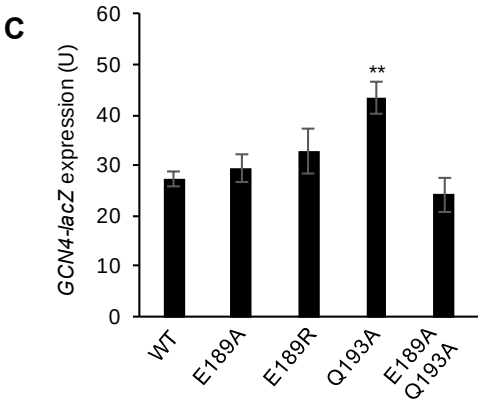
**Figure 3.10.8. Genetic and biochemical evidence that eIF2β substitutions at the eIF2β:Met-tRNA$_i$ interface of the open complex decrease the rate of TC binding and enhance the closed/P$_{IN}$ conformation at UUG codons**

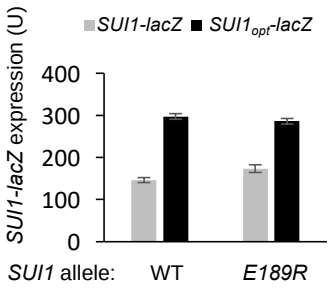
**(A)** Contacts between eIF2β and the Met-tRNA$_i$ ASL in the py48S-open complex. In py48S-closed eIF2β has moved away and instead contacts the Met-tRNA$_i$ D-loop (Fig. 7.2.1 B). eIF2β is shown in red, Met-tRNA$_i$ in green. Measurements are in angstroms. **(B)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring *HIS4-lacZ* reporters with AUG or UUG start codons were assayed for β-galactosidase activities as in Fig. 3.10.2 D. The ratio of expression of the UUG to AUG reporter was calculated from at least four different measurements, and the mean and S.E.M.s were plotted. **(C)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring plasmids (pPMB24 or pPMB25) with *SUI1-lacZ* fusions with either the native suboptimal ($_{-3}$CGU$_{-1}$) or optimized ($_{-3}$AAA$_{-1}$) AUG context were cultured and assayed for β-galactosidase activities as in Fig. 3.10.2 D. Mean expression levels and SEMs were calculated from four transformants, and relative (Rel.) mean expression levels normalized to that of the WT strain are listed, along with expression ratios for the *SUI1-lacZ* versus *SUI1$_{opt}$-lacZ* reporters. **(D)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring a *GCN4-lacZ* reporter on plasmid p180 were cultured in synthetic complete medium lacking leucine and uracil (SC-L-U) at 30°C to A$_{600}$ of ~0.8, and β-galactosidase activities were measured as in Fig. 3.10.2 D. Mean *GCN4-lacZ* expression (± SEM) in units of β-galactosidase activity was calculated from four transformants. **(E, F)** The $k_{on}$ values for TC binding to partial 43S·mRNA(AUG) or 43S·mRNA(UUG) complexes were determined as in Fig. 3.10.6 A-C for WT eIF2 and eIF2 containing eIF2β-S202A/K214A

from three independent experiments for each eIF2/mRNA combination. **(G-I)** TC dissociation rates were measured as in Fig. 3.10.5 D-G for WT eIF2 or eIF2 containing eIF2β-S202A/K214A for partial 43S·mRNA(AUG) or 43S·mRNA(AUG) complexes. One of three replicate experiments is shown for each eIF2/mRNA combination (G), from which mean rate constants (H) and end points (I) with SEMs were calculated. **(B-I)** Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (**, $p<0.01$).

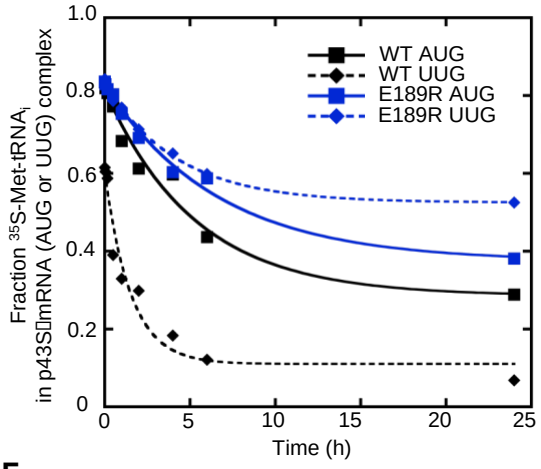



| *SUI1* allele: | WT | *E189R* |
|---|---|---|
| Rel. *SUI1-lacZ*: | 1.0 | 1.2 |
| Rel. *SUI1$_{opt}$-lacZ*: | 1.0 | 1.0 |
| *SUI1$_{opt}$-lacZ*/ *SUI1-lacZ*: | 2.0±0.09 | 1.6±0.09 |

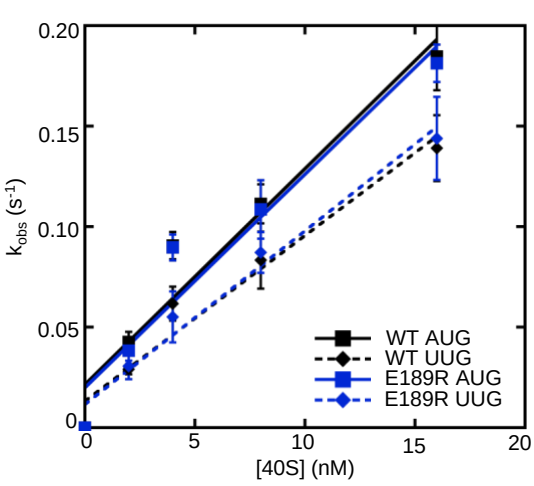



F

| | $k_{off}$ ($h^{-1}$) | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.66±0.12 | 0.17±0.02 |
| E189R | 0.27±0.05** | 0.19±0.02 |

G

| | End-point | |
|---|---|---|
| | mRNA | |
| eIF2β | UUG | AUG |
| WT | 0.09±0.02 | 0.25±0.04 |
| E189R | 0.53±0.01** | 0.38±0.01* |




I

| | $k_{on}$ ($s^{-1}$) | |
|---|---|---|
| | mRNA | |
| eIF2β | AUG | UUG |
| WT | 10.73±1.0 | 8.23±1.2 |
| E189R | 10.61±0.7 | 8.63±1.2 |

Fig. 9

**Figure 3.10.9. Genetic and biochemical evidence that eIF2β substitutions designed to alleviate a predicted clash between eIF2β and Met-tRNA$_i$ enhance transition to the closed state at UUG codons but do not alter the rate of TC loading**

**(A)** Overlaying the structure of eIF2β in py48S-open with Met-tRNA$_i$ in py48S-closed reveals predicted clashes between eIF2β (including residues E189 and Q193) and Met-tRNA$_i$ (including residue U45) that should impede rearrangement of Met-tRNA$_i$ to the closed state before AUG recognition. eIF2β is shown in red, Met-tRNA$_i$ in green. **(B)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring *HIS4-lacZ* reporters were analyzed as in Fig. 3.10.2 D. The ratio of expression of the UUG to AUG reporter was calculated from at least four independent transformants, and the mean and S.E.M.s were plotted. **(C)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring a *GCN4-lacZ* reporter were analyzed as in Fig. 3.10.2 D. Mean *GCN4-lacZ* (± SEM) was determined from at least four independent transformants. **(D)** Derivatives of strain KAY18 containing the indicated *SUI3* alleles and harboring plasmids (pPMB24 or pPMB25) with *SUI1-lacZ* fusions with either native suboptimal ($_{-3}$CGU$_{-1}$) or optimized ($_{-3}$AAA$_{-1}$) AUG context were analyzed as in Fig. 3.10.2 D. Mean expression levels and SEMs were calculated from four transformants, and relative (Rel.) mean expression levels normalized to that of the WT strain are listed, along with expression ratios for the *SUI1-lacZ* versus *SUI1$_{opt}$-lacZ* reporters. **(E-G)** TC dissociation rates were measured as in Fig. 3.10.5 D-G. One of three replicate experiments is shown for WT eIF2 and eIF2 containing the variant eIF2β-E189R (E), from which mean rate constants (F) and end points (G) with SEMs were calculated. **(H, I)** TC association rates were measured as in Fig. 3.10.6 A-C for WT eIF2 and eIF2 containing eIF2β-E189R and partial

43S·mRNA(AUG) or 43S·mRNA(UUG) complexes. The $k_{obs}$ values are plotted against the concentration of 40S subunits (H), and the calculated TC on-rates ($k_{on}$) are shown (I). Values are the averages of three independent experiments. **(B-I)** Asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t-test (**, $p<0.01$).

# 4. Ribosomal Protein Rps26/eS26 Plays Important Roles in Accurate Start Codon Selection for Translation Initiation

Laura Marler, Jyothsna Visweswaraiah, and Alan G. Hinnebusch†

Division of Molecular and Cellular Biology, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, Bethesda, Maryland 20892

†To whom correspondence should be addressed. Tel: +1-301-496-4480; Email: ahinnebusch@nih.gov

## 4.1 ABSTRACT

Translation initiation is the critical step of protein synthesis at which the correct start codon is selected, determining the sequence and quantity of proteins produced. In eukaryotes, a 43S preinitiation complex (PIC) consisting of the 40S ribosomal subunit, eukaryotic initiation factors (eIFs), and a ternary complex (TC) made up of GTP-bound eIF2 and methionyl initiator tRNA ($tRNA_i$) scans the mRNA leader for an AUG codon. Canonically, initiation occurs at the first AUG encountered, but it is known that the PIC may bypass an AUG in unfavorable sequence context. However, it is not known how the ribosome recognizes favorable context. Recent work from our lab indicates that ribosomal proteins, particularly the β–hairpin of Rps5/uS7, are involved in this process, and structural analyses of a 43S·mRNA complex reveal that Rps26/eS26 contacts the mRNA upstream of the AUG codon in the E site and mRNA exit channel of the 40S subunit near the Rps5/uS7 β-hairpin. To investigate the role of Rps26/eS26 in start codon selection, we mutated residues positioned near the mRNA and assayed for decreased initiation fidelity indicated by increased initiation at a near cognate UUG codon ($Sui^-$ phenotype) or suppression of UUG initiation in cells containing a $Sui^-$ mutation in another factor ($Ssu^-$ phenotype). We found that substitutions of residues of Rps26/eS26 that contact the mRNA further upstream (-10, -11) in the exit channel produced strong $Ssu^-$ phenotypes, increasing the requirement for an AUG and suppressing a $Sui^-$ mutation in another factor. These mutants are analogous to $Ssu^-$ mutations observed in Rps3/uS3 in the mRNA entry channel, suggesting a novel role for ribosomal proteins in clamping mRNA in both the entry and exit channels in order to maintain initiation fidelity. In addition, we discovered that Rps26/eS26 substitutions near the important -3, -4, and -5

nucleotides (numbered from the A of AUG (+1)) confer slow growth and decrease initiation fidelity (Sui$^-$ phenotype), indicating that these residues play a role in AUG selection. Because this region of Rps26/eS26 interacts with the mRNA backbone, it is possible that these residues help to position the mRNA for inspection by another factor, such as eIF2α or Rps5/uS7. Taken together, these data indicate important, previously unrecognized roles for Rps26/eS26 in context recognition and efficient and accurate start codon selection *in vivo*.

## 4.2 INTRODUCTION

Accurate start codon selection during translation initiation is critical to ensure that the correct cellular proteins are synthesized. In eukaryotes, this process generally occurs via a scanning mechanism. The small 40S subunit of the ribosome, along with eukaryotic initiation factors (eIFs) 1, 1A, 3, and 5, recruits a ternary complex (TC) consisting of methionyl initiator tRNA ($tRNA_i$) and GTP-bound eIF2 to form a 43S preinitiation complex (PIC). The PIC is recruited to the 5' end of mRNA, assisted by the eIF4F complex, and scans the 5'-untranslated region (UTR) for an AUG in favorable sequence context. The -3 and +4 positions (numbered from A of AUG (+1)) are of particular importance to sequence context optimality. During scanning, the 40S subunit assumes an open conformation in which $tRNA_i$ is not tightly bound in the peptidyl (P) site ($P_{OUT}$), and while hydrolysis of GTP may occur, the inorganic phosphate ($P_i$) is not released. eIF1 promotes the open conformation of the PIC by preventing full accommodation and tighter binding of $tRNA_i$ in the P site and by blocking $P_i$ release from eIF2-GDP·$P_i$. As discussed above, eIF2β is also involved in impeding transition to the closed state by stabilizing eIF1 binding and through its own interactions with $tRNA_i$. Upon recognition of an appropriate start codon, transition to the closed state of the PIC alters eIF2β interactions with eIF1 and $tRNA_i$ and eIF1 is released from the PIC, allowing $P_i$ release and full accommodation of $tRNA_i$ in the P site ($P_{IN}$). Subsequent dissociation of the other initiation factors is followed by joining of the large (60S) ribosomal subunit, catalyzed by eIF5B, resulting in an 80S initiation complex with $tRNA_i$ base-paired to AUG in the P site, ready to proceed with protein synthesis (reviewed in (12)).

As outlined above, eIF1 promotes the open, scanning-conducive conformation of the PIC to which TC rapidly loads in the $P_{OUT}$ state (151). eIF1 also blocks transition to the closed state at non-AUG codons or AUG codons in poor context (151, 209). This dual role of eIF1 positions it as a 'gatekeeper' of translation initiation fidelity, whose release from the PIC is required for start codon selection (129, 137, 207). Accordingly, mutations that weaken eIF1 binding to the 40S subunit reduce the rate of TC loading, derepressing expression of a *GCN4-lacZ* reporter ($Gcd^-$ phenotype) (212, 213). Such mutations also elevate initiation at near-cognate codons or AUGs in poor context ($Sui^-$ phenoypte) by causing premature eIF1 release, destabilizing the open PIC and favoring inappropriate transition to the closed/$P_{IN}$ state (213). Conversely, mutations that strengthen eIF1's interaction with the 40S subunit increase the stringency of the PIC for an AUG in good context ($Ssu^-$ phenotype) (214). Moreover, decreasing the cellular abundance of eIF1 decreases initiation accuracy, while overexpression of eIF1 suppresses initiation at near-cognate or poor context start codons (146, 196, 212, 215, 350). This link between eIF1 abundance and initiation accuracy is exploited to autoregulate eIF1 expression, since the eIF1 gene (*SUI1* in yeast) initiates from an AUG in poor context and the frequency of its recognition is inversely related to eIF1 abundance (215).

The optimal sequence context around an AUG (summarized in Fig. 4.6.1 A) plays a large role in start codon selection (351). In most vertebrates, the Kozak consensus sequence of gccRccAUGG is preferred, where upper case letters represent less variability than lowercase ones and R represents a purine (A or G) (304, 306). The presence of a purine at the -3 position is the most highly conserved element of this sequence (after AUG), with

97% of vertebrate mRNAs displaying a purine at this position and A being the more common nucleotide (305). The introduction of point mutations at different positions of the consensus sequence confirmed the importance of a purine, preferably A, at the -3 position (307). In yeast, the optimal start codon context is poly(A), conserving the preference for a purine at the -3 position, with 81% of all yeast mRNAs and 100% of highly expressed yeast genes containing an A at -3 (309). Despite the demonstrated importance of sequence context to start codon selection, and therefore, to gene expression, how the ribosome recognizes the context nucleotides is not well understood at the molecular level.

During scanning, the mRNA enters the 40S subunit through a 12 nucleotide entry channel, threads around the neck through the A, P, and E decoding sites, and exits through a 12 nucleotide exit channel above the platform (Fig. 4.6.1 B) (91, 261, 352). As a result, when an AUG is positioned in the P site, where it is decoded by Met-tRNA$_i$, the -1 to -3 nucleotides lie in the E site, and context nucleotides beginning with the -4 position are located in the exit channel. The α-subunit of eIF2 lies in the vicinity of the important -3 nucleotide (109, 195) and crosslinking studies indicate that the N-terminal domain (D1) of eIF2α contacts the mRNA at the -3 position in the exit channel (265). In addition, eIF2βγ complexes lacking eIF2α were shown to form 48S complexes less efficiently than heterotrimeric eIF2 complexes, and the absence of eIF2α reduced the stimulatory effect of a purine at the -3 position on PIC assembly in a reconstituted mammalian system (265). These observations indicate that eIF2α-D1 may be involved in recognition of the -3 nucleotide during start codon selection.

Structural studies have revealed that the conserved β-hairpin of ribosomal protein Rps5/uS7 lies in the mRNA exit channel near eIF2α-D1 and the -3 nucleotide of mRNA (Fig. 4.6.2) (92, 195). Recent studies from our lab have demonstrated that the β-hairpin of Rps5/uS7 is required for efficient and accurate start codon selection *in vivo* and *in vitro* (314). Substitutions in the β-strand 1 of this hairpin reduce the efficiency of AUG recognition and exacerbate the effect of poor context at the *SUI1* start codon, resulting in decreased cellular abundance of eIF1 and an indirect increase in initiation at near-cognate UUG codons that can be rescued by restoring WT levels of eIF1. Meanwhile, substitutions in the loop portion of the Rps5 β-hairpin discriminate against near-cognates by destabilizing the $P_{IN}$ state specifically at non-AUG triplets. Thus, overall, substitutions in the Rps5 β-hairpin destabilize the $P_{IN}$ state and decrease initiation at suboptimal initiation codons. In addition, differential contacts at the Rps5:eIF2α interface between the open and closed states of the PIC were shown to be instrumental in stabilizing each state respectively, modulating the transition to $P_{IN}$ and promoting accurate AUG selection (316). These findings underscore the importance of Rps5/uS7 for start codon recognition and demonstrate the active role that ribosomal proteins near the decoding sites and exit channel can play in AUG selection and context discrimination.

Four ribosomal proteins have loops or tails that extend into the mRNA exit channel (92, 195). In addition to the β-hairpin of Rps5/uS7 discussed above, these include Rps14/uS11, Rps26/eS26, and Rps28/eS28 (Fig. 4.6.2). Aside from Rps5/uS7, only Rps26/eS26 is predicted to directly contact the important -3 context nucleotide. In

comparing the contacts of mRNA in the bacterial and eukaryotic ribosomes, the most striking differences were found in the region 5' of the decoding sites, with Rps26/eS26, for which there is no bacterial counterpart, comprising the main component of the divergent site (353). There is evidence that Rps26 is required to maintain the preference for the optimal Kozak sequence in yeast (318). In addition, structural data indicates that a well-conserved α-helix of Rps26/eS26 contacts the mRNA backbone at the -3, -4, and -5 nucleotides, opposite the β-hairpin of Rps5/uS7 (Fig. 4.6.3 A-B) (195). As such, Rps26/eS26 is not likely to be involved in directly sensing nucleotide identity in this region. Rather, we hypothesized that Rps26/eS26 may play an important role in positioning mRNA to allow nucleotide recognition by some other component of the PIC (possibly eIF2α or Rps5/uS7). To date, no single residue or component has been identified, the elimination of which resulted in a complete loss of context recognition. It is possible that a base stacking interaction between the G903 of yeast 18S rRNA and the -3 nucleotide could play such a pivotal role, but it is increasingly likely that context recognition is mediated through a combination of factors, or possibly by modulation of the shape of the mRNA channel itself. In this scenario, the preference for a purine at the -3 position could be dictated by the size of the nucleotide base rather than its specific molecular contacts, and Rps26/eS26 could play an important role in restricting the size of the channel.

Additionally, crosslinking studies have implicated well conserved residues of Rps26/eS26 in interactions with mRNA further upstream in the exit channel, at the -4 to -9 nucleotides (Fig. 4.6.4 A-C) (353). The well conserved residues K66, Y68, and K70

were all seen to form crosslinks to this region (317). Recent cryo-EM structures of the ribosome that resolve mRNA nucleotides up to the -10 position indicate that the upstream interaction with Rps26 is likely to begin at the -10 nucleotide, which contacts Rps26-K66, and extend further towards the mRNA 5' end (111). It has been shown by our lab that substitutions of residues in ribosomal protein Rps3/uS3 that contact mRNA in the entry channel destabilize the closed PIC, producing the hyperaccuracy phenotypes of reduced UUG initiation in a Sui$^-$ mutant and decreased initiation at the poor-context AUG of *SUI1* (eIF1) mRNA (119). This suggests a role for ribosomal proteins in the entry channel in stabilizing the closed state on start codon recognition through interactions with mRNA, possibly assisting the cessation of scanning. We hypothesized that, by its interactions with mRNA in the exit channel, Rps26/eS26 plays a role in stabilizing the closed state analogous to that of Rps3/uS3 in the entry channel.

In this study, we have investigated the role of Rps26/eS26 in both context recognition and start codon selection *in vivo*. We find that substitutions of Rps26/eS26 residues K66, Y68, and K70 that contact mRNA upstream of the -10 nucleotide confer strong hyperaccuracy (Ssu$^-$) phenotypes. These substitutions increase the stringency of AUG selection, decreasing initiation at UUG codons in the presence of the strong Sui- mutant eIF5-G31R (encoded by the *SUI5* allele of *TIF5*). Mutations that destabilize the closed complex typically discriminate against not only near cognate start codons, but also AUG codons in poor sequence context, and we find that indeed, these substitutions result in decreased recognition of poor-context start codons. Thus, our data are consistent with the conclusion that these highly conserved residues of Rps26/eS26 that contact mRNA in the

exit channel are involved in stabilizing the closed PIC. Additionally, we report that substituting the residues of Rps26/eS26 near the -3 nucleotide, particularly H80, results in weak hypoaccuracy ($Sui^-$) phenotypes. Frequently, mutations conferring $Sui^-$ phenotypes, relaxing the requirement for an AUG start codon, also relax the requirement for optimal context for AUG codons (200, 212-214, 302). However, our Rps26 substitutions have effects on context recognition that do not always follow this convention, apparently decoupling start codon selection and context recognition. Thus, our data demonstrate a novel and important role for Rps26/eS26 in accurate and efficient start codon recognition through stabilizing the closed PIC conformation, and further suggest a role in context recognition through a distinct mechanism.

## 4.3 MATERIALS AND METHODS

### 4.3.1 Plasmid and yeast strain constructions

Yeast strains used in this study are listed in Table 4.6.1. The strain JVY05 was generated from H2994 (*MATa ura3-52 trp1-Δ63 leu2-3 leu2-112 his4-303(AUU)*) by the one-step PCR strategy (354), using the *hphMX6* cassette to replace *RPS26B* and selecting for resistance to hygromycin on rich medium containing galactose as a carbon source (YPGal). The $P_{GAL1}$-*RPS26A* strain JVY09 was similarly generated from JVY05 by the one-step PCR strategy (354) to insert the $P_{GAL1}$ promoter immediately upstream of the RPS26A ORF using the *kanMX6* cassette and selecting for resistance to kanamycin on YPGal. Replacement of *RPS26B* with the *hphMX6* cassette and integration of the *kanMX::*$P_{GAL1}$ promoter cassette at *RPS26A* were verified by PCR analyses of genomic

DNA using the appropriate primers. JVY09 was shown to be inviable on glucose medium (where the *GAL1* promoter is repressed) in a manner fully complemented by plasmid-borne *RPS26A* on pJVB06 (Fig. 4.6.6). Derivatives of JVY09 harboring high copy *LEU2* plasmids containing WT (pJVB06) or mutant *RPS26A* alleles (pLMP5 – pLMP23), listed in Table 4.6.2, were generated by transformation on a rich galactose medium lacking leucine. Derivatives of JVY09 harboring low copy *LEU2* plasmids containing WT (pJVB05) or mutant *RPS26A* alleles (pLMP75 – pLMP79) were generated similarly.

Plasmids used in this study are listed in Table 4.6.2. The low-copy *LEU2* plasmid pJVB05 was made by inserting into pRS315 (355) a 2.3 kb BamHI restriction fragment containing *RPS26A* flanked by 1120 bp upstream and 832 bp downstream of the coding sequences, amplified from genomic DNA of strain H2994. The insert from pJVB05 was sub-cloned into pRS425 (p1398) to create the high-copy *LEU2* plasmid pJVB06. To create plasmids pLMP5 – pLMP23, mutations were introduced into *RPS26A* in pJVB06 using the QuikChange site-directed mutagenesis system (Stratagene) according to the manufacturer's instructions and using the primers listed in Table 4.6.3. pLMP75 – pLMP79 were generated in a similar manner by performing site-directed mutagenesis on pJVB05. All constructs were verified by DNA sequencing of the entire *RPS26A* ORF.

### 4.3.2 Biochemical assays using yeast cell extracts

Assays of β-galactosidase activity in whole cell extracts (WCEs) were performed as described previously (356). The sequence context of the start codon for AUG and UUG *HIS4-lacZ* reporters is 5'-AUA(AUG/UUG)G-3'. For Western analyses, WCEs were

prepared by trichloroacetic acid (TCA) extraction as previously described (345) and immunoblot analysis was conducted as described in (210) with antibodies against eIF1/Sui1 and Hcr1(128). Enhanced chemiluminescence (Amersham) was used to visualize immune complexes, and signal intensities were quantified by densitometry using NIH ImageJ software.

### 4.3.3 Polysome profile analysis

For polysome analysis, strains were grown in SD + His + Ura + Trp at 30° C to $A_{600}$ of 0.8 – 1.0. Cyclohexamide was added (50 μg/ml) 5 min prior to harvesting, and WCEs were prepared in breaking buffer (20 mM Tris-HCL, pH 7.5, 50 mM KCl, 10 mM MgCl2, 1 mM dithiothreitol, 5 mM NaF, 1 mM phenylmethylsulfonyl fluoride, 1 Complete EDTA-free Protease Inhibitor Tablet (Roche)/ 50 ml buffer). Fifteen $A_{260}$ units of each WCE were separated by velocity sedimentation through a 4.5-45% sucrose gradient by centrifugation at 39,000 rpm for 3 hours in an SW41Ti rotor (Beckman). Gradient fractions were scanned at 254 nm to visualize ribosomal species.

## 4.4 RESULTS

### 4.4.1 Structural analysis of Rps26/eS26 contacts with mRNA

Crosslinking evidence indicates that a region of Rps26/eS26 consisting of residues 61 – 70 contacts the mRNA upstream from the AUG codon, in the exit channel (261, 265, 317, 353). This region of Rps26 contains five residues that are highly

conserved throughout eukaryotes (Fig. 4.6.4 A). Notably, a positive charge at position 66 (usually a lysine), a tyrosine at position 68, and a lysine at position 70 are completely conserved, identifying these residues as strong candidates for interaction with the mRNA. Available crystal structures did not resolve the mRNA in this region (a structure of a partial mammalian 48S PIC (pm48S) containing mRNA, $tRNA_i$, and eIF1A resolves mRNA only to the -6 nucleotide (195)), but a more recent partial yeast 48S PIC in the closed conformation (py48S-closed) assembled with TC, eIF1, eIF1A, and mRNA visualized via cryo-EM, resolves mRNA to the -10 nucleotide and indicates that this region of Rps26/eS26 is likely to contact the mRNA from the -10 to -12 nucleotides (Fig. 4.6.4 B-C) (111). One contact between the Rps26/eS26 K66 and the -10 nucleotide can be seen in this latter structure. Conserved basic residues in Rps3/uS3 in the entry channel were found to be important for the stability of the closed PIC (119). Thus, we hypothesized that Rps26/eS26 contacts with mRNA in this upstream region would impact start codon selection.

Analysis of the pm48S structure reveals that residues 73 – 81 of Rps26/eS26 form an α-helix that lies near the mRNA upstream of the decoding sites in the vicinity of the important -3, -4, and -5 context nucleotides (Fig. 4.6.3 A) (195). This helix is well conserved in evolution, and a histidine at Rps26/eS26 position 80, the only residue to directly contact the -3 nucleotide, is conserved throughout eukaryotes, as well as a positively charged residue at position 82 (K in yeast) (Fig. 4.6.3 B). Our detailed analysis of these structures revealed multiple contacts between this helix and mRNA. Rps26-His80 sits between the -3 and -4 nucleotides and forms multiple close (< 3 Å) contacts

with the mRNA backbone. A serine at position 81 (alanine in *S. cerevisiae*) occupies a similar position between the -4 and -5 nucleotides. Together, these two residues appear to form a three-dimensional platform for the mRNA backbone that could potentially hold it in position for optimal interactions with PIC components on the face of the mRNA containing the nucleotide bases (Fig. 4.6.5 A). Analysis of the cryo-EM structure of a yeast partial 48S PIC containing TC, eIF1, eIF1A, eIF3, and mRNA in the closed state (py48S-closed) reveals similar contacts (Fig. 4.6.5 B) (111).

In addition to its interactions with the mRNA, the Rps26/eS26 α-helix described above also contacts rRNA nucleotides in helix 28 (h28). Helix 28 forms a "neck" between the head and body of the 40S subunit, and is compressed in the closed state, but relaxed in the open PIC (111). Mutations altering the A928:U1389 base pair or the bulge G926 in this helix confer $Gcd^-$ phenotypes that are suppressed by overexpression of $tRNA_i^{Met}$, indicating that they are important for the rate or stability of TC binding to the PIC (197). h28 is involved in head rotation during the transition of the PIC from the open to closed state upon start codon recognition, and a change in the pitch of h28, along with changes in the orientation of eIF2β, drives the closure of the mRNA entry channel latch and helps to arrest scanning (111). It is therefore possible that contacts between Rps26/eS26 and h28 might modulate the rigidity of that helix, impacting start codon selection. A lysine at position 82 in Rps26/eS26, just outside of the α-helix, contacts both the mRNA and rRNA helix 28. This positively charged residue, conserved in higher eukaryotes, lies between the mRNA and helix 28, contacting the -5 context nucleotide

(Fig. 4.6.5 C). We hypothesized that this residue could play a role in context recognition and modulate helix 28 rigidity.

In addition to its α-helix, a loop of Rps26/eS26 also contacts the mRNA at the important context nucleotides near the start codon. Residue N25 in this loop contacts the -4 and -5 nucleotides, making contacts with atoms of both the mRNA backbone and the -4 nucleotide (Fig. 4.6.5 A-B). This position might allow N25 to sample nucleotide identity in addition to aiding H80 and other α-helix residues in positioning mRNA.

These contacts of the Rps26/eS26 α-helix, K82, and N25 with mRNA near the -3 context nucleotide are corroborated by the more recent py48S-closed structure (Fig. 4.6.5 B) (111). This structure also supports our interpretation of the role of Rps26-H80, which is seen to contact the mRNA backbone between the -3 and -4 nucleotides in a very similar position to that seen in the pm48S PIC. While a serine at position 81 is well conserved throughout higher eukaryotes, the yeast Rps26/eS26 contains an alanine at this position, which assumes a slightly different position from that of the mammalian structure. It is therefore possible that H80 alone fulfills this function in yeast, while higher eukaryotes have evolved a longer platform for mRNA positioning in this region.

### 4.4.2 Substitutions in Rps26/eS26 residues that contact mRNA in the exit channel increase discrimination against near-cognate UUG codons *in vivo*

Residues 60 - 71 in the β-sheet region of Rps26/eS26 were implicated in crosslinking to mRNA at the -6 to -10 nucleotides (317). The recent py48S-closed

structure that resolves mRNA structure to the -10 nucleotide indicates a possible interaction between the -10 position and Rps26-K66 (111). Based on this structure, interactions between Rps26/eS26 and bases further upstream appear likely (Fig. 4.6.4 C). To investigate this possible interaction, we generated yeast strains containing Rps26/eS26 variants substituted for the conserved residues K66, Y68, and K70. Each residue was substituted with Ala, to shorten the side chain, or Glu, to introduce a negative charge. *RPS26A* alleles encoding these variants on hc plasmids were introduced into a strain (JVY09) in which *RPS26B* was deleted and chromosomal *RPS26A* was placed under the inducible galactose promoter, $P_{GAL1}$ (Fig. 4.6.6 A). We demonstrated that neither *RPS26B* deletion nor replacing the native *RPS26A* promotor with $P_{GAL1}$ alters the ability of the strain to grow on galactose medium compared to the WT parent strain (H2994) (Fig. 4.6.6 B, sectors 2 and 4 compared to sector 1). Because expression of $P_{GAL1}$-*RPS26A* in strain JVY09 is repressed by glucose, no WT Rps26/eS26 can be produced when cells are grown on glucose-containing medium, rendering JVY09 cells inviable (Fig. 4.6.6 B, sector 9). The introduction of WT *RPS26A* on a plasmid restores Rps26 expression on glucose medium, complementing the lethal phenotype (Fig. 4.6.6 B, sector 10).

JVY09 was transformed with high copy plasmids containing each *RPS26A* allele and mutant phenotypes were scored following a switch from galactose to glucose medium (SD +Ura +His), where $P_{GAL1}$-*RPS26* expression is repressed. These strains contain a *his4-303* allele in which the AUG start codon of *HIS4* has been mutated to AUU, preventing initiation at the canonical start codon and conferring histidine auxotrophy. The resulting $His^-$ phenotype (Fig. 4.6.7, row ii) is suppressed by $Sui^-$

mutations, which decrease the stringency of start codon selection, allowing initiation at the third, in-frame UUG codon, restoring expression of the gene product of *HIS4* and allowing histidine biosynthesis (Fig. 4.6.7, row iii). This Sui$^-$/His$^+$ phenotype was observed in JVY09 on introduction of the known dominant *SUI5* allele of *TIF5*, encoding eIF5-G31R (235) (Fig. 4.6.9, row 2 vs. 1). A mutation that increases the stringency of start codon selection will suppress a His$^+$, Sui$^-$ phenotype imparted by a mutation in another factor, preventing initiation at UUG and rendering a strain His$^-$ (Fig. 4.6.7, row iv).

We found that substitutions of the Rps26/eS26 residues of interest produced Slg$^-$ phenotypes, particularly those that introduce a charge switch (*K66E* and *K70E*) (Fig. 4.6.8, left panel, rows 6 and 9 compared to row 1). They did not suppress the His$^-$ phenotype of *his4-303* however (Fig. 4.6.8, right panel), indicating that none of them confer obvious Sui$^-$ phenotypes. We next tested the mutant alleles for the ability to suppress the elevated UUG initiation at *his4-303* and attendant His$^+$ phenotype, as well as the Slg$^-$ phenotype on medium containing histidine, conferred by the dominant Sui$^-$ mutation *SUI5* (Fig. 4.6.9, row 2 vs. row 1). We found that substitutions in this region, particularly those with acidic side chains, suppressed both the His$^+$ and Slg$^-$ phenotypes of *SUI5*. In the presence of *SUI5*, *K66E* and *Y68A* eliminated both the His$^+$ and Slg$^-$ phenotypes (Fig. 4.6.9, rows 5 and 6), and *K70E* significantly diminished growth on medium lacking histidine (Fig. 4.6.9, row 8). Thus, it appears that these mutations impair recognition of the UUG start codon of *his4-303* mRNA, conferring an Ssu$^-$ phenotype.

Mutations that increase the stringency of start codon selection typically also suppress the elevated UUG:AUG initiation ratio, measured using paired *HIS4-lacZ* reporters with either an AUG or a UUG start codon, which is increased by ~5-fold over WT in the strain harboring *SUI5* and WT *RPS26A* (Fig. 4.6.10, cols. 1-2). We found that all three of the *rps26* substitutions that suppressed the $His^+$ phenotype of *SUI5* also significantly reduced the *HIS4-lacZ* UUG:AUG ratio. *K66E*, *Y68A*, and *K70E* each approximately halved the UUG:AUG ratio in the presence of *SUI5*, reducing initiation at UUG codons from 18% observed with WT *RPS26A* to 8%, 10%, and 9%, respectively (Fig. 4.6.10, rows 3-5 vs. 1). These results indicate that substitutions of residues in the Rps26/eS26 β-sheet region restore discrimination against near-cognate UUG codons in the presence of *SUI5*, increasing the fidelity of start codon selection.

### 4.4.3 Substitutions in Rps26/eS26 residues that contact the mRNA in the exit channel discriminate against initiation codons in suboptimal sequence context

Other $Ssu^-$ mutations have been found to discriminate against poor-context AUG codons to a degree that parallels their discrimination against non-AUG start codons (314). We therefore asked whether substitutions in the β-sheet region of Rps26/eS26, which display strong $Ssu^-$ phenotypes *in vivo*, also discriminate against AUG codons in suboptimal context. To investigate this possibility, we tested whether the substitutions reduce recognition of the AUG codon of a uORF, and thereby increase expression of the downstream ORF encoded on the same mRNA, by the process of leaky scanning. For this purpose, we employed *GCN4-lacZ* reporters in which *GCN4* uORF1 is elongated to overlap the *GCN4* CDS start codon (el.uORF1) and the WT optimal context of the AUG

codon of uORF1 ($^{-3}$AAA$^{-1}$) has either been maintained or mutated to weak ($^{-3}$UAA$^{-1}$) or poor ($^{-3}$UUU$^{-1}$) context (Fig. 4.6.11 A). With optimal context at el.uORF1, virtually all scanning ribosomes recognize the uORF1 AUG and subsequent reinitiation at the *GCN4* ORF following el.uORF1 is very rare, resulting in extremely low *GCN4-lacZ* expression (Grant, 1994). Replacing the WT context with weak context at uAUG-1, by introducing a U at the -3 position, increases leaky scanning of el.uORF1 and elevates *GCN4-lacZ* expression by ~6 fold. When the optimal context is instead replaced with the poor, poly-(U) context, leaky scanning is further elevated and an ~30 fold increase in *GCN4-lacZ* expression is observed (Fig. 4.6.11 B). To estimate the maximum possible initiation at the *GCN4* ORF, we also measured expression of a construct in which uAUG-1 was eliminated entirely, resulting in an ~100-fold increase in *GCN4-lacZ* expression. Based on these results, we were able to calculate the percentage of scanning ribosomes that initiate at el.uORF1 (Fig. 4.6.11 C), and the percentage that bypass (leaky scan) the upstream ORF to initiate at the *GCN4* ORF (Fig. 4.6.11 D). In WT cells, the percentages of scanning ribosomes that reach the *GCN4* AUG and translate *GCN4-lacZ* are ~1%, ~6%, and ~30% for optimum, weak, and poor contexts, respectively (Fig. 4.6.11 D). Therefore, ~99%, ~94%, and ~70% of scanning ribosomes recognize uAUG-1 in optimum, weak, and poor contexts, respectively (Fig. 4.6.11 C). Note that although *GCN4-lacZ* expression increases ~30 fold with poor context at el.uORF1, this represents only an ~30% reduction in recognition of uAUG-1, since virtually no leaky scanning (~1%) occurs when uAUG-1 is in optimum context. Note also that the small difference in deduced uORF1 translation between weak and optimum context of ~94% and ~99% (Fig.

4.6.11 C) is meaningful because it is based on the highly significant differences in *GCN4-lacZ* expression for these two constructs (Fig. 4.6.11 B).

Assaying expression of these *el.uORF1-GCN4-lacZ* reporters revealed that the Ssu$^-$ substitutions in Rps26/eS26 moderately decrease initiation at uAUG-1 in weak and poor contexts. All of the mutants increased leaky scanning of uAUG-1, seen as an increase of *GCN4-lacZ* expression (Fig. 4.6.12, red and green bars for mutants versus WT, Table 4.6.4). The strongest effect was seen for the *K70E* substitution, which displayed deduced levels of ~72% and ~40% initiation at the el.uORF1 in weak or poor context, respectively, compared to the corresponding values of ~94% and ~70% for WT ribosomes (Fig. 4.6.12 C, red and green bars, col. 5 vs. col. 1; Table 4.6.4 C, column 5 vs. 1). These results are consistent with an increase in the stringency of the requirement for optimal start codon context, which often accompanies the increased stringency of start codon selection seen for these mutants.

We next tested whether substitution of Rps26/eS26 residues in the β-sheet resulted in lowered eIF1 levels. Substitutions that discriminate against AUG codons in poor context also decrease eIF1 levels, since the AUG of the *SUI1* gene is itself in suboptimal context. We found that all of the Rps26/eS26 Ssu$^-$ mutants decreased eIF1 levels to between 65% and 22% of that of WT cells, with stronger effects resulting from the introduction of an acidic residue (Fig. 4.6.13). Thus, it appears that removing a positive charge (K66 or K70) or a bulky side chain (Y68), and especially introducing a negative charge in this region, disrupts the interaction of Rps26/eS26 with mRNA,

resulting in a more stringent requirement for optimal sequence context for AUG selection. This is consistent with the hypothesis that these substitutions destabilize the closed PIC, thus requiring greater stabilization from a perfect codon:anticodon match or optimal context nucleotides for the ribosome to overcome this disruption and enter the closed state.

### 4.4.4 Effects of Rps26/eS26 substitutions on ribosome biogenesis are not likely to cause the phenotypes seen

In eukaryotes, the 40S and 60S ribosomal subunits are assembled from rRNAs and ribosomal proteins in the nucleus before being exported to the cytoplasm for final processing. A defect in 40S biogenesis results in an excess of 60S subunits that are unable to form 80S complexes due to a lack of available 40S subunits, lowering the 40S:60S ratio, as seen in mutations of the conserved 40S ribosome biogenesis factor, Ltv1 (357). In addition, general translation is impaired when 40S subunits are not readily available, resulting in a decrease in polysomes (elongating ribosomes). Because mutations in ribosomal proteins have the potential to disrupt ribosome biogenesis, we tested whether these substitutions affected the ratio of 40S to 60S subunits. Plasmids bearing WT or mutant *RPS26A* were introduced into JVY09 and the resulting strains were grown to log phase in media containing glucose. Whole cell extracts obtained from these strains were submitted to ultracentrifugation through a sucrose gradient (4.5% – 45%). Velocity sedimentation in this manner paired with UV spectroscopy allows the resolution of cell extracts into the following peaks: light fractions, free 40S subunits, free 60S subunits, 80S complexes, and polysomes.

We found that Rps26A-K66E displayed a similar free 40S:60S ratio relative to cells containing WT RPS26A on a high copy plasmid (Fig. 4.6.14 A). However, we found that introduction of the WT hc plasmid produced a low 40S:60S ratio compared to the parent strain (Fig. 4.6.14 B, iii vs. i), which was not observed for WT Rps26/eS26 on a lc plasmid (Fig. 4.6.14 B, ii vs. iii). Hence, we generated the mutants *K66E*, *Y68A*, and *K70E* in a lc plasmid and introduced them into the same parent strain. In the presence of *his4-303*, low copy plasmids containing *K66E* and *Y68A* suppressed the $Slg^-$ and $His^+$ phenotypes of *SUI5*, and *K70E* suppressed the $His^+$ phenotype, similar to our observations of these mutants in the high copy plasmid (Fig. 4.6.15, rows 4, 5, and 7 vs. 2). In polysome profiles, we found that these substitutions reduced the free 40S:60S ratio compared to lc WT *RPS26A* (Fig. 4.6.15, ii – iv vs. i). The decrease in 40S:60S ratio is comparable to, although greater than that seen for the WT hc plasmid (Fig. 4.6.14 B iii vs. Fig. 4.6.15 ii – iv). Thus, it is likely that both these substitutions and the WT hc plasmid result in a 40S biogenesis defect. We cannot rule out that a possible 40S biogenesis defect could affect the phenotypes demonstrated above. However, we note that the hc WT Rps26/eS26 plasmid does not confer a detectable $Sui^-$ phenotype (Fig. 4.6.8, row 1 vs. row 4; Fig. 4.6.17, row 1) or $Ssu^-$ phenotype (Fig. 4.6.9, row 2; Fig. 4.6.20, row 2) despite its affect on free 40S:60S ratio. We also observed an effect on the polysome:monosome ratio for all three mutants, indicating the possibility of a general defect in translation initiation.

### 4.4.5 Substitutions in Rps26/eS26 residues that contact the -3, -4, and -5 context nucleotides impair start codon selection

In order to examine the role of contacts between Rps26/eS26 and mRNA context nucleotides in start codon selection, we introduced substitutions into Rps26/eS26 residue N25 and residues in the α-helix (I79, H80, and A81) that contact the -3, -4, and -5 nucleotides, as well as residue K82 that contacts the -6 nucleotide and h28 of the 18S rRNA (Figs. 4.6.3 and 4.6.5). These residues were generally substituted with Ala to shorten the side-chain, or with basic or acidic residues to introduce or alter side-chain charge. *RPS26A* alleles encoding these variants were introduced on hc plasmids into the strain described above (JVY09) in which *RPS26B* was deleted and chromosomal *RPS26A* was placed under the inducible galactose promoter ($P_{GAL1}$) (Fig. 4.6.6 A). Despite strong sequence conservation, substitution of residues in the α-helix of Rps26/eS26 generally did not affect yeast growth, although the charge switch for variant H80E resulted in a slow growth phenotype (Fig. 4.6.17 A, row 6 vs. 1). While R82E and R82Q did not confer significant growth phenotypes (Fig. 4.6.17 A, rows 13 – 14 vs. 1), the R82A substitution was lethal (not shown), suggesting that a bulky residue, rather than a specific charge, may be required at position 82. The substitution I79D was also lethal (not shown). In the py48S-closed structure, the sidechain of I79 is oriented inward toward the center of the α-helix, suggesting that inserting a negative charge in this position destabilizes the helix (111).

None of the introduced Rps26/eS26 substitutions was found to suppress the $His^-$ phenotype of *his4-303* (Fig. 4.6.17 B, rows 5 - 14), indicating that they do not confer

obvious Sui$^-$ phenotypes. However, a known Sui$^-$ allele of *SUI3*, encoding eIF2β-S264Y (*SUI3-2*), also did not suppress the His$^-$ phenotype in JVY09 (Fig. 4.6.17 B, row 3; Fig. 4.6.17, sector 2). Because of this last observation, we screened for synthetic His$^+$ phenotypes for the Rps26/eS26 variants in the presence of *SUI3-2*. We found that Rps26/eS26 variants H80A, N25A/H80A, N25D, R82E, R82Q, and H80E all produced synthetic His$^+$ phenotypes to different degrees, suggesting that these variants confer a range of weak Sui$^-$ phenotypes (Fig. 4.6.18, sectors 3 - 8). Of the single mutations tested, *H80E* produced the strongest synthetic His$^+$ phenotype (Fig. 4.6.18, sector 8), while the double mutation *N25A/H80A* conferred the strongest His$^+$ phenotype overall (Fig. 4.6.18, sector 4).

The inability of these Rps26/eS26 substitutions to confer a His$^+$ phenotype on their own suggests that they do not elevate the UUG:AUG ratio sufficiently to produce enough *his4-303* product for adequate histidine biosynthesis (358). This has been seen previously for the eIF1 mutation *sui1-K37A* and Rps5/uS7 substitutions *E144R* and *R225K*, which were shown to elevate UUG initiation moderately without suppressing the His$^-$ phenotype of *his4-303* (213, 314). In order to explore this possibility, we assayed the expression of the matched *HIS4-lacZ* reporters containing either an AUG or UUG start codon. Moderate increases in the UUG:AUG initiation ratio were seen for seven of the Rps26 variants, including *H80E* (3.4-fold), *A81E* (2.8-fold), *R82Q* (2.1-fold), *R82E* (1.9-fold), *H80A* (1.8-fold), *N25D* (1.7-fold), and *I79K* (1.6-fold) (Fig. 4.6.19). Surprisingly, the double substitution *N25A/H80A*, displayed a significantly lower UUG:AUG ratio than that of WT *RPS26A* (0.7-fold), which is discussed further below. Previous findings

suggest that a $His^+$ phenotype in a *his4-303* strain is associated with an increase in the UUG:AUG initiation ratio of > 3.5-fold (213). Thus, the moderate increases in the UUG:AUG ratios and lack of $His^+$ phenotypes is consistent with the idea that most substitutions in the α-helix region result in weak $Sui^-$ phenotypes, suggesting a limited role for this segment of Rps26/eS26 in stringent AUG selection.

### 4.4.6 *N25A/H80A* and *A81D* increase the stringency of start codon selection

As discussed above, the Rps26-N25A/H80A variant conferred a synthetic $His^+$ phenotype in the presence of *SUI3-2* in a *his4-303* strain, indicative of a weak $Sui^-$ phenotype, but reduced rather than increased the UUG:AUG ratio in otherwise WT cells. Previously, it was observed that certain eIF5 substitutions confer a $Sui^-$ phenotype in otherwise WT cells, but are able to suppress the $Sui^-$ phenotype conferred by the *SUI5* mutation in eIF5 (236). Although the molecular basis for this complex genetic interaction is unknown, it prompted us to determine whether the N25A/H80A variant is capable of suppressing *SUI5*. Interestingly, both N25A/H80A and A81D suppress the $His^+$ phenotype of *SUI5* and N25A/H80A additionally suppresses the $Slg^-$ phenotype conferred by *SUI5* (Fig. 4.6.20, rows 6 and 8). None of the other Rps26/eS26 substitutions in this region affected growth in a *SUI5* strain. Co-suppression of both $Slg^-$ and $His^+/Sui^-$ phenotypes of *SUI5* was described previously for $Ssu^-$ substitutions in eIF1 that restore initiation accuracy and reduce UUG initiation in *SUI5* cells (214).

In summary, a group of substitutions in this region of Rps26/eS26, exemplified by the H80E variant, appear to moderately reduce the accuracy of start codon selection, increasing the UUG:AUG initiation ratio in otherwise WT cells and conferring a $His^+$ phenotype in *SUI3-2* cells. By contrast, the N25A/H80A variant appears to have the opposite effect on accuracy, reducing the UUG:AUG ratio in otherwise WT cells and suppressing the $Sui^-$ and $Slg^-$ phenotypes of *SUI5*. In future work, it will be interesting to determine whether the N25A/H80A mutation suppresses the elevated UUG:AUG initiation ratio conferred by *SUI5*. It is difficult to explain the synthetic $His^+$ phenotype conferred by N25A/H80A in *SUI3-2* cells. This variant appears to alter start codon selection by a complex mechanism that is influenced in opposite ways by different $Sui^-$ mutations in eIF5 versus eIF2β. Further study of N25A/H80A may shed light on the differences between the mechanisms of eIF5-G31R and eIF2β-S264Y as well as the role of Rps26/eS26 in context recognition.

### 4.4.7 Substitutions in Rps26/eS26 residues near the -3 nucleotide have nuanced effects on context recognition

Because residues of Rps26/eS26 in the α-helix contact the mRNA at the important -3 nucleotide and upstream context nucleotides, we hypothesized that substitutions of these residues might influence the effect of context on selection of AUG start codons. Mutants that confer $Sui^-$ phenotypes often do so by destabilizing the open state of the PIC and favoring transition to the closed state, increasing the recognition of suboptimal start codons, whether near cognate (e.g., UUG) start codons or AUGs in poor context. For example, substitutions in eIF1 that reduce its interaction with eIF2β in the

open PIC, thus destabilizing eIF1 binding, not only increase UUG initiation but also increase initiation at *GCN4* uAUG-1 in poor context (302). Thus, we anticipated that mutations in the α-helix of Rps26/eS26 that confer weak Sui$^-$ phenotypes might decrease discrimination against weak and poor context to a modest degree. At odds with this expectation, however, these mutations resulted in decreased initiation at uAUG-1 in weak and poor context, suggesting increased, rather than relaxed discrimination against non-optimal sequence context. *I79K*, *H80A*, *H80E*, and *A81D*, all of which conferred Sui$^-$ phenotypes of varying strength, displayed increased leaky scanning and attendant increased *GCN4-lacZ* expression in the el.-uORF1 reporters containing weak or poor context (Fig 4.6.21; Table 4.6.5, columns 2 – 4, 6, 7, and 9).

These unanticipated results might point to a unique mechanism by which substitution of these residues perturbs context recognition. Because they are positioned near and, in some cases, in direct contact with the -3 nucleotide, it is possible that these residues alter context recognition in a manner unrelated to altered stability of the open or closed complex, and consequently, their effects on AUG context are decoupled from their reduced discrimination against non-AUG start codons. In general, these substitutions result in moderate effects on weak context. *I79K*, *A81K*, *A81D*, and *R82E* reduce the percentage of ribosomes initiating at el.uORF1 in weak context from ~94% in WT to 80 – 87 % (Table 4.6.5 C, columns 2, 6, 7, and 9; row 2). The same substitutions confer relatively greater effects on the recognition of poor context, reducing initiation from ~70% in WT to between ~30% and ~59% (Table 4.6.5 C, columns 2, 6, 7, and 9 vs. 1; row 3). Substitutions of Rps26/eS26 H80 resulted in larger increases in leaky scanning.

These mutants decrease initiation at uAUG-1 even in optimum context, indicating that they confer a general initiation defect, in addition to stronger discrimination against AUG codons in weak and poor context. For each reporter, the introduction of a negative charge at position 80 confers a stronger defect than merely removing the histidine side chain. At uAUG-1 in optimum context, *H80A* and *H80E* reduce the WT initiation rate of ~99% to ~95% and ~90%, respectively (Table 4.6.5 C, columns 3 and 4; row 1). In weak context, they dramatically lower recognition of uAUG-1 from the WT ~94% to ~56% and ~36%, respectively (Table 4.6.5 C, columns 3 and 4; row 2), while their respective initiation rates at uAUG-1 in poor context are ~29% and ~14%, 2.4-fold and 5-fold reductions compared to 70% for WT ribosomes (Table 4.6.5 C, columns 3 and 4; row 3).

We observed similar increases in leaky scanning at weak and poor contexts for the Ssu$^-$mutants *A81D* and *N25A/H80A*. Although this result parallels the discrimination against near cognate codons seen in these mutants, we cannot rule out that they also discriminate against poor context through a unique mechanism as a result of their direct contacts with the context nucleotides. Indeed, they confer very similar context effects to those seen in the Sui$^-$ substitutions in this region. *A81D* results in ~98%, ~80%, and ~42% initiation at uAUG-1 in optimum, weak, and poor contexts, respectively (Table 4.6.5 C, column 7). *N25A/H80A* had a more pronounced effect, reducing initiation to ~97%, ~64%, and ~20% in optimum, weak, and poor contexts, respectively (Table 4.6.5 C, column 5). Only *A81F* was found to reduce the stringency of context recognition, increasing initiation at uAUG-1 to ~99%, ~98%, and ~89% in optimum, weak, and poor contexts, respectively (Table 4.6.5 C, column 8). Thus, we see that substitutions in the α-

helix of Rps26/eS26 have distinctive effects on context recognition that do not always correlate with their impacts on initiation at near cognate codons, suggesting that they may impact context recognition through a novel mechanism not directly tied to start codon selection.

## 4.5 DISCUSSION

In this study, we obtained genetic evidence implicating the 40S ribosomal protein Rps26/eS26 in ensuring efficient and accurate start codon recognition *in vivo*. In recent crystal and cryo-EM structures of the ribosome (111, 195), Rps26/eS26 approaches the key context nucleotide at the -3 position of mRNA, and projects into the mRNA exit channel in a manner predicted to contact the mRNA context upstream of the -10 nucleotide. Evidence from our lab implicated Rps5/uS7, another ribosomal protein in the mRNA exit channel, in altering start codon recognition through its interactions with eIF2$\alpha$, which also contacts the -3 nucleotide (109, 110, 314, 316).

In the recent py48S-closed structure (111), mRNA in the exit channel was visualized to the -10 nucleotide, which contacts Rps26/eS26 residue K66. We infer from this structure that other well conserved, basic residues in Rps26/eS26 implicated in cross-linking to mRNA (Y68, K70) lie in proximity to nucleotides upstream of the -10 position. We found that substitutions of these residues suppress recognition of a UUG start codon in *HIS4* mRNA resulting from a known strong Sui$^-$ variant of eIF5 (*SUI5*), thus conferring Ssu$^-$ phenotypes. Initiation at UUG codons was more dramatically suppressed when a basic residue was replaced with an acidic one (K66E, K70E). Substitutions of

K66, Y68, and K70 were also found to reduce expression of eIF1, indicating increased discrimination against the poor sequence context of the *SUI1* start codon. Accordingly, K66E, Y68A, and K70E decrease initiation at the AUG codon of *GCN4* uORF1 when it resides in poor context. Thus, substitutions of these residues increase discrimination against not only non-AUG start codons, but also AUG start codons in poor sequence context.

Hyperaccuracy phenotypes such as those seen for substitutions of K66, Y68, and K70 are typical of mutations that destabilize the closed PIC (212, 214, 316), indicating that residues in this region of Rps26/eS26 may play a role in stabilizing the closed state of the ribosome following AUG recognition. This phenomenon has also been seen in basic residues of Rps3/uS3 that contact mRNA in the entry channel. Dong, et al. found that substitution of R116 or R117 of Rps3/uS3 to acidic residues suppressed utilization of UUG start codons and discriminated against the native, poor context of the eIF1 AUG start codon (119). These substitutions were also found to increase the dissociation rate of TC *in vitro*, supporting the conclusion that the observed Ssu$^-$ phenotypes result from destabilization of the $P_{IN}$ state, particularly in the presence of the inherently less stable codon:anticodon duplex formed at a UUG start codon. Further, Rps3-R116D and –R117D were found to destabilize 48S PICs assembled on model mRNAs designed to leave the mRNA exit channel empty or only partially occupied (119). *In vitro* studies of TC binding kinetics and the dependence of complex stability on entry channel contacts in the presence of Rps26/eS26 mutants K66E, Y68A, and K70E in the future would be informative. We hypothesize that, like Rps3/uS3 in the entry channel, Rps26/eS26 may

function in the exit channel to stabilize mRNA binding in the closed state of the ribosome, aiding in the cessation of scanning following start codon selection.

Our genetic findings indicate that substitutions in the Rps26/eS26 α-helix near the -3 nucleotide impact start codon selection and context recognition, suggesting this region plays a role in ensuring the accuracy of these processes. These findings are summarized in Fig. 4.6.21. The residues of this α-helix that lie nearest to the mRNA appear to act as a platform, cradling the mRNA backbone. Substitutions of H80, which lies between the -3 and -4 nucleotides, to Ala or Glu resulted in Sui$^-$ phenotypes, with stronger phenotypes resulting from the acidic side chain, while substitutions of A81, which lies between the -4 and -5 nucleotides, to basic Arg or acidic Asp produced opposing phenotypes (Sui$^-$ for *A81K*, Ssu$^-$ for *A81D*). Substitution of R82 to Ala was lethal, while substitution with acidic side chains (Asp or Glu) produced Sui$^-$ phenotypes. Additionally, substitutions of N25 produce Sui$^-$ phenotypes while the double mutant N25A/H80A confers an Ssu$^-$ phenotype. If the effects of Rps26/eS26 mutants on start codon selection were the result of altered scanning speed, we would expect to see the introduction of a positive charge result in more interaction with the negatively charged mRNA, slowing scanning. Decreased scanning speed would result in longer dwell times over non-optimal start codons, allowing initiation at near cognate codons, thus conferring a Sui$^-$ phenotype. Introduced negative charges, which would repel the mRNA, would be expected to increase scanning speed and result in an Ssu$^-$ phenotype. However, this relationship is not seen. Rather, it appears that the effects of Rps26/eS26 residues in this region are nuanced and may be highly dependent on the shape of the mRNA exit channel. Additionally,

because these residues contact the mRNA backbone, they may function to position mRNA nucleotides for specific contacts on the nucleotide face, potentially with eIF2α or Rps5/uS7. Thus, it appears that Rps26, along with other components of the exit channel, creates a contoured platform for mRNA in order to maintain the optimum position for AUG recognition and stability of the codon:anticodon duplex.

Mutations in the initiation factors and other components of the PIC that affect start codon selection typically have analogous effects on context recognition. For example, mutations in eIF1 that alter eIF1 binding to the 40S subunit produce strong Sui$^-$ phenotypes by causing premature eIF1 release, reducing discrimination against near cognate start sites such as UUG (213). For the same reason, these mutants also reduce discrimination against AUG codons in poor context, resulting in increased eIF1 levels (213). There have been many instances in which a Sui$^-$ phenotype coincides with increased recognition of poor context, while Ssu$^-$ mutations are commonly seen to coincide with increased discrimination against poor context (200, 302, 316). These concurrent phenotypes are typically presumed to result from the same mechanism (e.g., eIF1 release, open state destabilization, or altered scanning speeds). However, in the case of Rps26/eS26 mutants, this expected relationship is not seen. Instead, all but one mutant (A81F) result in increased discrimination against AUG codons in poor context, regardless of their effects on start codon recognition. Decoupling of start codon recognition from context recognition may be a result of the proximity of these residues to the -3, -4, and -5 context nucleotides, allowing them to act on these two processes via different mechanisms. It is possible that contacts between mRNA and components of the exit

channel at the nucleotide face normally act to slow scanning to ensure appropriate start codon selection. Perturbations in Rps26/eS26 that disrupt such contacts may increase scanning speed, resulting in increased discrimination against poor context, while changing the shape of the exit channel, thus altering the stability of the codon:anticodon duplex to produce a Sui$^-$ phenotype. Substitutions in other ribosomal proteins of residues that contact mRNA have been observed to affect start codon recognition without affecting context recognition (119, 314). While the data presented here do not implicate the mechanism by which these substitutions alter the recognition of start sites or context, it is clear that this region of Rps26/eS26 plays an important and nuanced role in start codon selection.

Extensive research into the contacts of initiation factors and ribosomal proteins that contact the mRNA at the -3 nucleotide has yet to conclusively implicate a single interaction in the recognition of this important context nucleotide (265, 314, 316). While some interactions remain to be explored (e.g., the basic C-terminal tails of Rps14/uS11 and Rps28/eS28, a potential stacking interaction with the G903 nucleotide of the 18S rRNA), we suggest that a system of interactions that rely on the shape of the mRNA exit channel function together to ensure accurate context recognition. Our data regarding the nuanced roles of Rps26/eS26 residues in context recognition, decoupled from start codon selection, supports this hypothesis. In addition, we find that residues of Rps26/eS26 contacting the mRNA at the -10 nucleotide and further upstream play an important role in stabilizing the closed PIC following AUG recognition. Thus, we describe novel and vital

roles for Rps26/eS26, a ribosomal protein in the mRNA exit channel in ensuring accurate start codon selection.

## 4.6 Figures and Tables

**A**

**Consensus sequence:**

**5' Cap**

**Mammals (Kozak)** $A/G^{-3}C^{-2}C^{-1}\underline{AUG}G^{+4}$

≥32 nts ***S. cerevisiae*** $A^{-3}A^{-2}A/G^{-1}\underline{AUG}A^{+4}$ **mRNA**

**B**

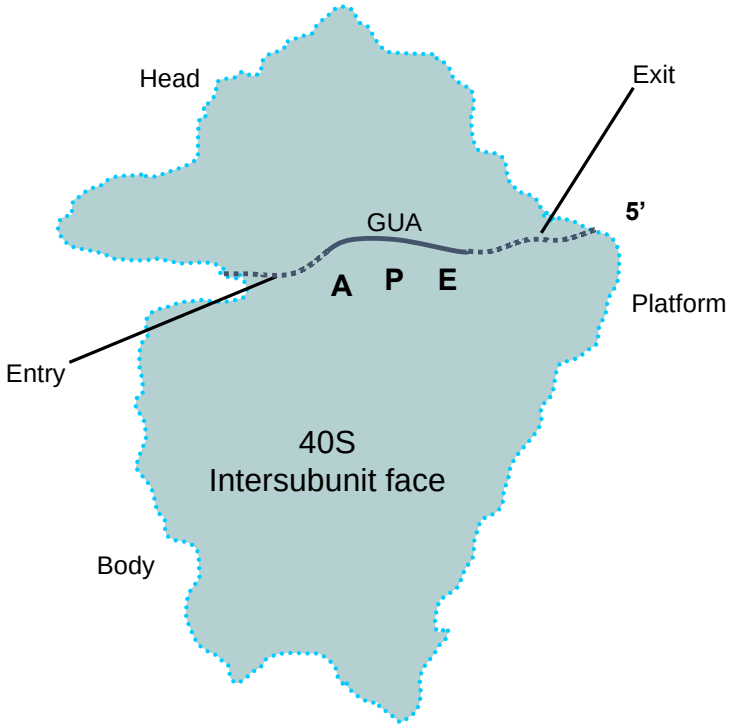


**Figure 4.6.1. Identity and position of context nucleotides affects start codon choice.**

**(A) Consensus sequences show optimal AUG context.** Consensus sequences for mammals (Kozak context, above) and S. cerevisiae (below) are shown. Nucleotides are numbered beginning with the A of AUG (+1). Genetic experiments have corroborated these sequences as optimal start codon context and revealed that the most important determinants of good context in yeast are the presence of an A in the -3 position and a purine at +4.

**(B) Path of mRNA through the 40S ribosomal subunit.** mRNA enters the 40S subunit through a 12 nucleotide entry channel above the shoulder, threads around the neck of the small subunit through the A, P, and E sites, and exits through a 12 nucleotide exit channel above the platform. As a result, when an AUG is in the P site, where it is decoded by initiator tRNA, the upstream context nucleotides -1 to -3 lie in the E site, and context nucleotides beginning with the -4 position are located in the exit channel.

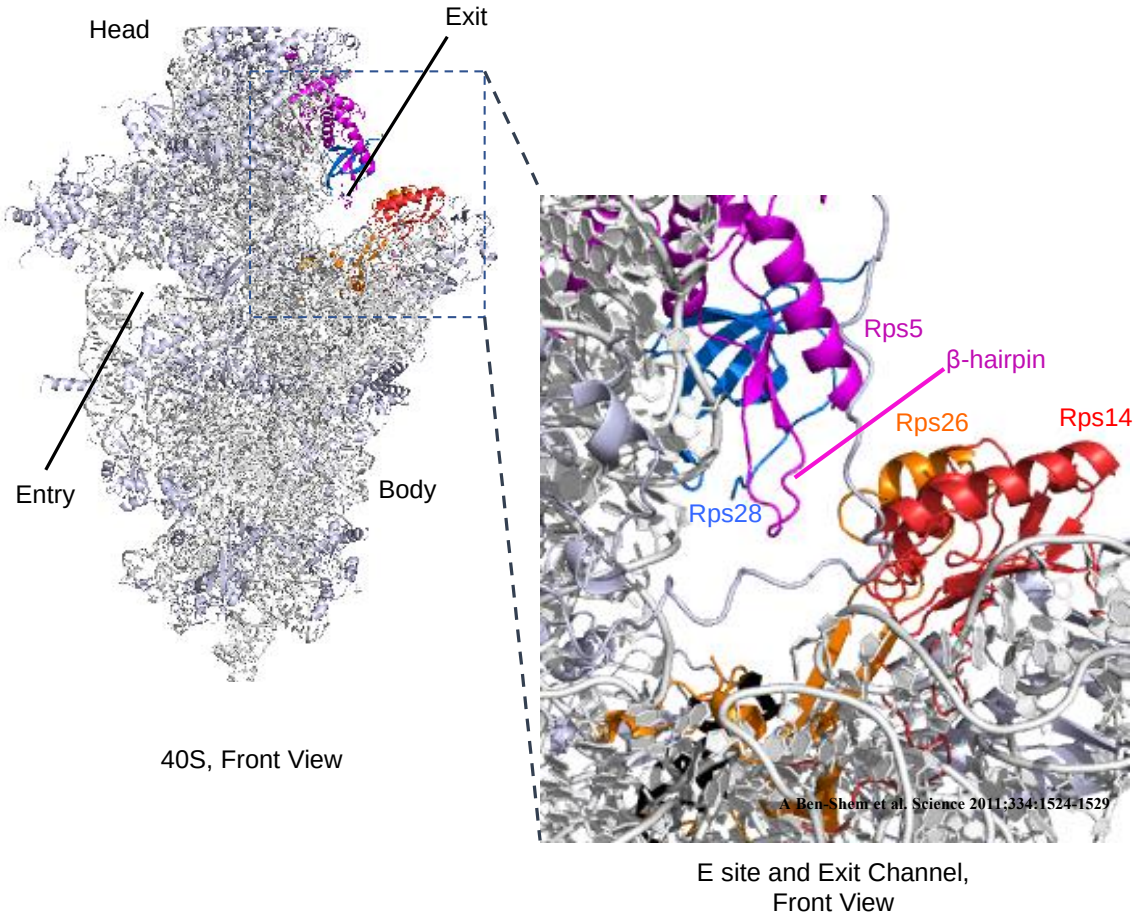



**Figure 4.6.2 Location of ribosomal proteins in the mRNA exit channel.**

Depiction of partial yeast 48S PIC (PDB 3J81) showing Rps5/uS7 (magenta), Rps14/uS11 (red), Rps26/eS26 (orange), and Rps28/eS28 (blue). For clarity, other ribosomal proteins are not shown. 18S rRNA in grey.

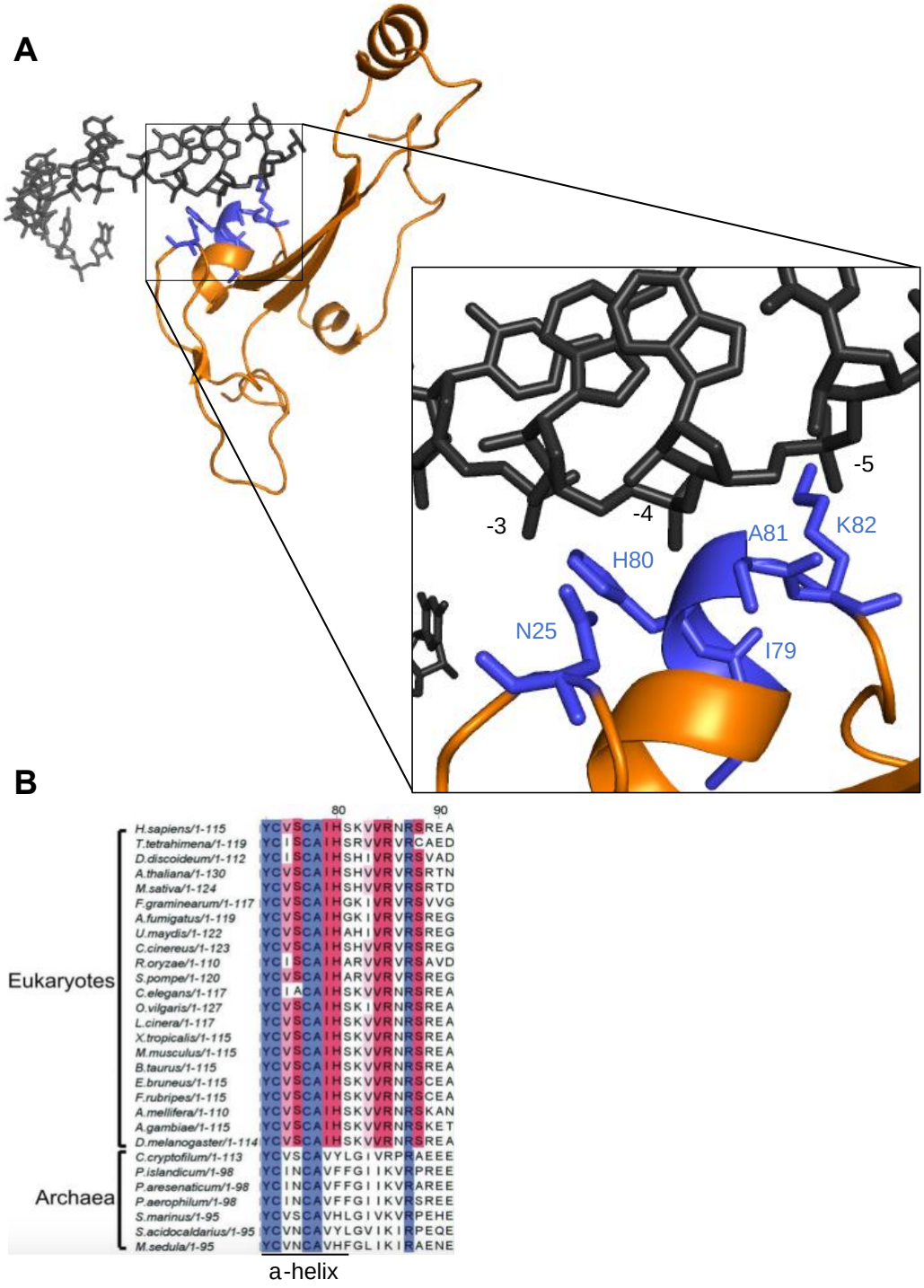

A
-5
-4
-3
A81
K82
H80
N25
I79
B
80
90
H.sapiens/1-115 YCVSCAIHSKVVRNRSREA
T.tetrahimena/1-119 YCISCAIHSRVVRVRCAED
D.discoideum/1-112 YCISCAIHSHIVRVRSVAD
A.thaliana/1-130 YCVSCAIHSHVVRVRSRTN
M.sativa/1-124 YCVSCAIHSHVVRVRSRTD
F.graminearum/1-117 YCVSCAIHGKIVRVRSVVG
A.fumigatus/1-119 YCVSCAIHGKIVRVRSREG
U.maydis/1-122 YCVSCAIHAHIVRVRSREG
C.cinereus/1-123 YCVSCAIHSHVVRVRSREG
R.oryzae/1-110 YCISCAIHARVVRVRSAVD
S.pompe/1-120 YCVSCAIHARVVRVRSREG
C.elegans/1-117 YCIACAIHSKVVRNRSREA
O.vilgaris/1-127 YCVSCAIHSKIVRNRSREA
L.cinera/1-117 YCVSCAIHSKVVRNRSREA
X.tropicalis/1-115 YCVSCAIHSKVVRNRSREA
M.musculus/1-115 YCVSCAIHSKVVRNRSREA
B.taurus/1-115 YCVSCAIHSKVVRNRSREA
E.bruneus/1-115 YCVSCAIHSKVVRNRSCEA
F.rubripes/1-115 YCVSCAIHSKVVRNRSCEA
A.mellifera/1-110 YCVSCAIHSKVVRNRSKAN
A.gambiae/1-115 YCVSCAIHSKVVRNRSKET
D.melanogaster/1-114 YCVSCAIHSKVVRNRSREA
C.cryptofilum/1-113 YCVSCAVYLGIVRPRAEEE
P.islandicum/1-98 YCINCAVFFGIIKVRPREE
P.aresenaticum/1-98 YCINCAVFFGIIKVRAREE
P.aerophilum/1-98 YCINCAVFFGIIKVRSREE
S.marinus/1-95 YCVSCAVHLGIVKVRPEHE
S.acidocaldarius/1-95 YCVNCAVYLGVIKIRPEQE
M.sedula/1-95 YCVNCAVHFGLIKIRAENE
Eukaryotes
Archaea
a-helix

**Figure 4.6.3. Residues of Rps26/eS26 α-helix contacting mRNA are well conserved.**

**(A)** Depiction of partial yeast 48S PIC (PDB 3JAP) showing Rps26/eS26 (orange) and mRNA (black). Residues that contact mRNA are shown in sticks and colored blue. For clarity, other ribosomal proteins and rRNA are not shown.

**(B)** Alignment of Rps26/eS26 sequences from diverse eukaryotes and archaea. Residues conserved in all kingdoms are marked in blue; residues conserved only in eukaryotes are in red. More intense color corresponds to a higher extent of conservation. Modified from Sharifulin, et al., 2012 under Creative Commons CC BY-NC 3.0.

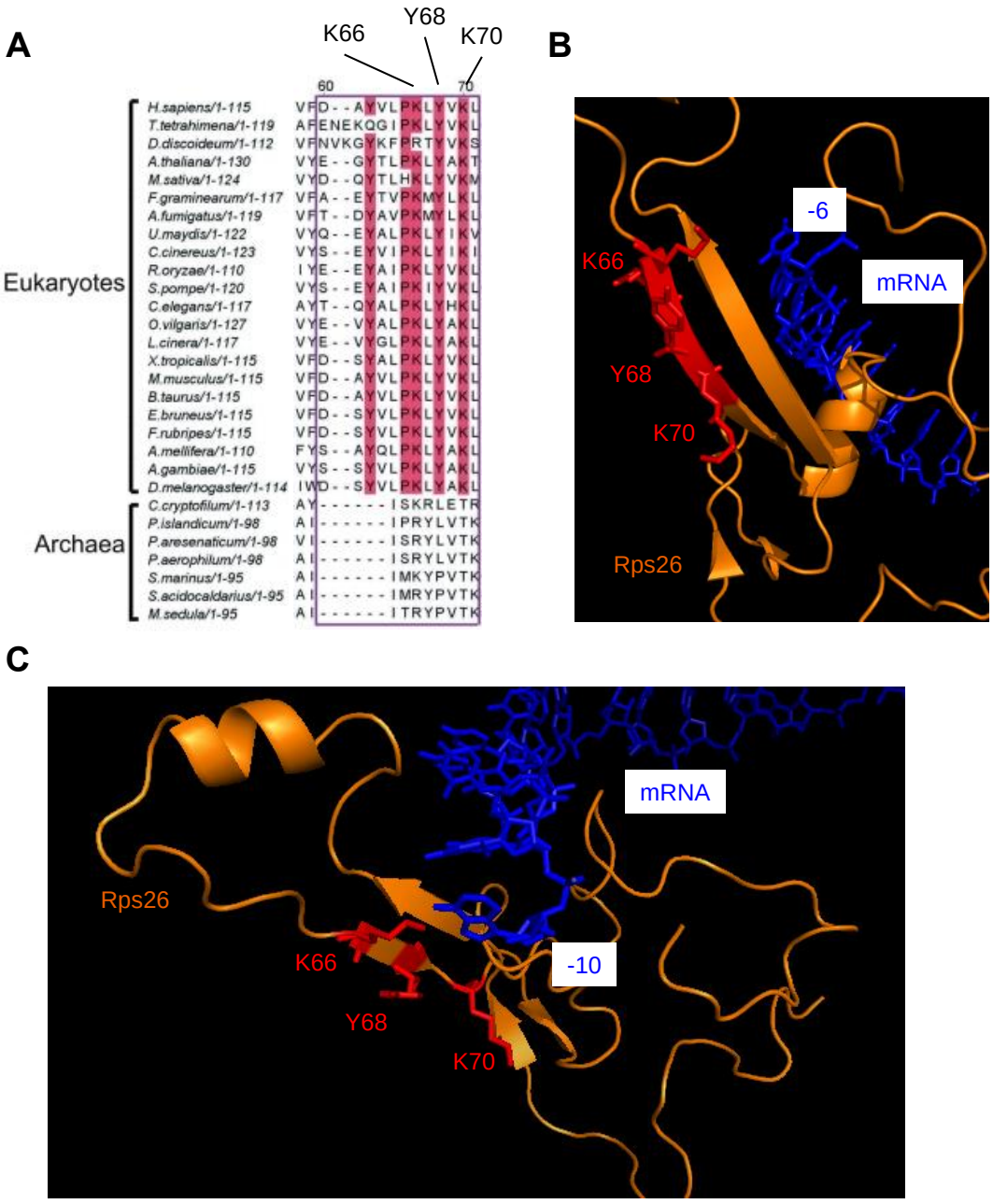
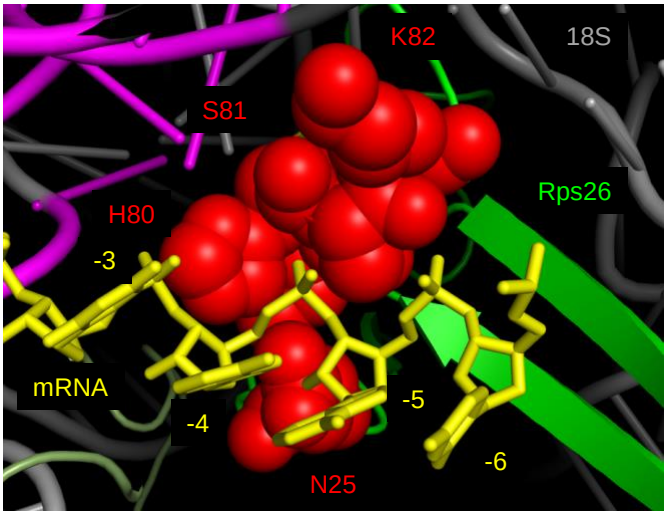

A
K66
Y68
K70
60
70
Eukaryotes
H.sapiens/1-115 VFD--AYVLPKLYVKL
T.tetrahimena/1-119 AFENEKQGIPKLYVKL
D.discoideum/1-112 VFNVKGYKFPRTYVKS
A.thaliana/1-130 VYE--GYTLPKLYAKT
M.sativa/1-124 VYD--QYTLHKLYVKM
F.graminearum/1-117 VFA--EYTVPKMYLKL
A.fumigatus/1-119 VFT--DYAVPKMYLKL
U.maydis/1-122 VYQ--EYALPKLYIKV
C.cinereus/1-123 VYS--EYVIPKLYIKI
R.oryzae/1-110 IYE--EYAIPKLYVKL
S.pompe/1-120 VYS--EYAIPKIYVKL
C.elegans/1-117 AYT--QYALPKLYHKL
O.vilgaris/1-127 VYE--VYALPKLYAKL
L.cinera/1-117 VYE--VYGLPKLYAKL
X.tropicalis/1-115 VFD--SYALPKLYVKL
M.musculus/1-115 VFD--AYVLPKLYVKL
B.taurus/1-115 VFD--AYVLPKLYVKL
E.bruneus/1-115 VFD--SYVLPKLYVKL
F.rubripes/1-115 VFD--SYVLPKLYVKL
A.mellifera/1-110 FYS--AYQLPKLYAKL
A.gambiae/1-115 VYS--SYVLPKLYAKL
D.melanogaster/1-114 IWD--SYVLPKLYAKL
Archaea
C.cryptofilum/1-113 AY------ISKRLETR
P.islandicum/1-98 AI------IPRYLVTK
P.aresenaticum/1-98 VI------ISRYLVTK
P.aerophilum/1-98 AI------ISRYLVTK
S.marinus/1-95 AI------IMKYPVTK
S.acidocaldarius/1-95 AI------IMRYPVTK
M.sedula/1-95 AI------ITRYPVTK
B
-6
mRNA
K66
Y68
K70
Rps26
C
mRNA
Rps26
K66
Y68
K70
-10

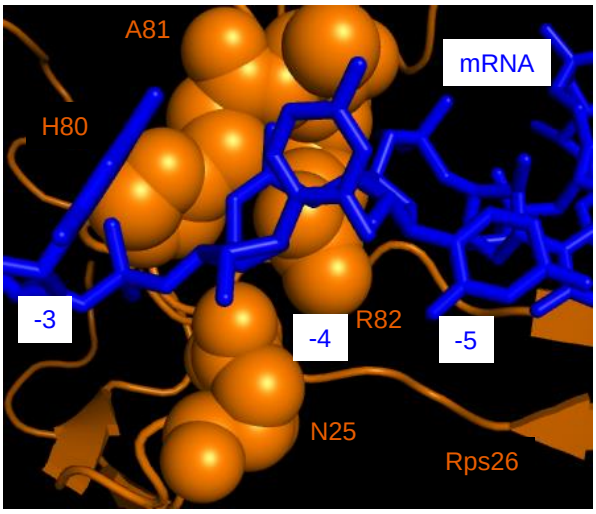
**Figure 4.6.4. Contacts of Rps26/eS26 with mRNA in the exit channel.**

**(A)** Alignment of Rps26/eS26 sequences from diverse eukaryotes and archaea. Residues conserved in all kingdoms are marked in blue; residues conserved only in eukaryotes are in red. The region implicated in crosslinking is boxed in purple. Modified from Sharifulin, et al., 2012 under Creative Commons CC BY-NC 3.0.
**(B)** Depiction of rabbit ribosome (PDB 4KZZ)showing Rps26/eS26 in orange and mRNA in blue. Residues to be modified are shown as sticks and colored red. **(C)** Depiction of partial yeast 48S PIC (PDB 3JAP) showing Rps26/eS26 in orange and mRNA in blue. Residues to be modified are shown as sticks and colored red.

A


K82
18S
S81
Rps26
H80
-3
mRNA
-4
-5
-6
N25

B


A81
mRNA
H80
-3
-4
R82
-5
N25
Rps26

C

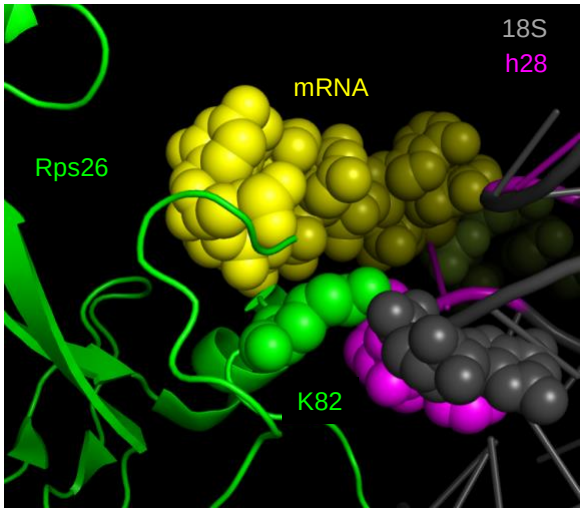

18S
h28
mRNA
Rps26
K82

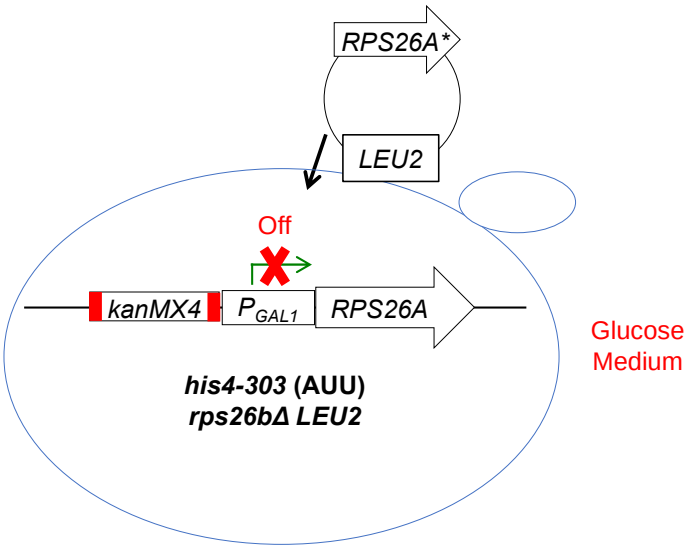
**Figure 4.6.5. Rps26/eS26 residues contact mRNA upstream of the start codon.**

**(A)** Depiction of rabbit ribosome (PDB 4KZZ) with Rps26/eS26 in green, residues to be modified shown in spheres and colored red, mRNA in yellow, and 18S rRNA in grey. Nucleotides of the rRNA that contact mRNA in the exit channel are colored in magenta. The depiction shows the possible role of Rps26/eS26 in acting as a platform to position the -3, -4, and -5 mRNA context nucleotides.

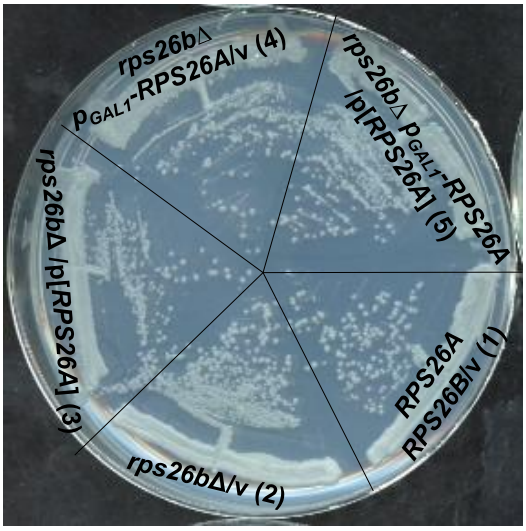
**(B)** Depiction of partial yeast 48S PIC (PDB 3JAP) showing Rps26/eS26 in orange and mRNA in blue. Residues in Rps26/eS26 to be modified are shown as spheres.

**(C)** Depiction of rabbit ribosome (PDB 4KZZ) with Rps26/eS26 in green, mRNA in yellow, and 18S rRNA in grey. Nucleotides of the rRNA that contact mRNA in the exit channel are colored in magenta. Rps26/eS26 residue K82 is shown in spheres to illustrate the position this residue occupies between the mRNA and h28 of 18S rRNA.

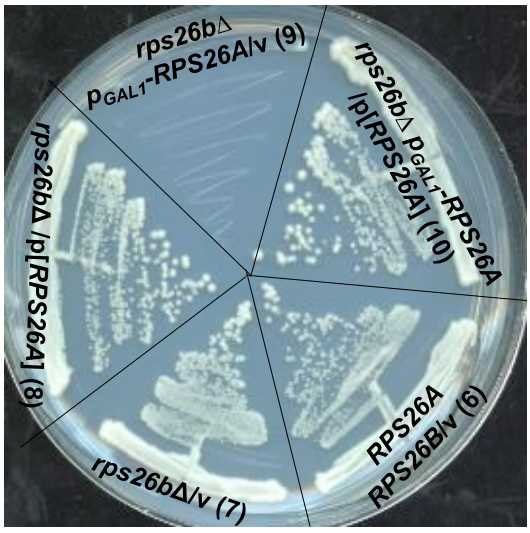
**A**




**B**




SGal -Leu




SC -Leu

**Figure 4.6.6. Introduction of *RPS26A* rescues growth of galactose inducible RPS26 strain on glucose.**

**(A)** Depiction of strain JVY09, in which *RPS26B* is deleted and *RPS26A* is placed under the inducible galactose promoter ($P_{GAL1}$). *RPS26A* variants were introduced into this strain by transformation on either high copy or low copy *LEU2* vectors. When grown on glucose medium, the chromosomal *RPS26A* gene is not expressed.

**(B)** Yeast cells from the parent strain H2994 containing a high copy LEU2 vector (pRS425) (*RPS26A RPS26B*/v), JVY05 containing pRS425 (*RPS26B*D/v), JVY05 containing RPS26A on a high-copy *LEU2* plasmid (pJVB06) (*rps26B*D/p[RPS26A]), JVY09 containing pRS425 (*RPS26B*D *pGAL-RPS26A*/v), and JVY09 transformed with a high copy *RPS26A LEU2* plasmid (pJVB06) (*RPS26B*D *pGAL-RPS26A*/p[*RPS26A*]) were streaked on synthetic complete galactose (SGal -Leu) or glucose (SC -Leu) media and grown at 30º C for 2 and 3 days, respectively. It can be seen that deletion of *RPS26B* does not significantly hamper growth, nor does the introduction of p[*RPS26A*]. Placing *RPS26A* under the $P_{GAL1}$ promoter in a strain deleted for *RPS26B* prevents growth on glucose medium, and the introduction of *RPS26A* on a plasmid rescues growth on glucose.

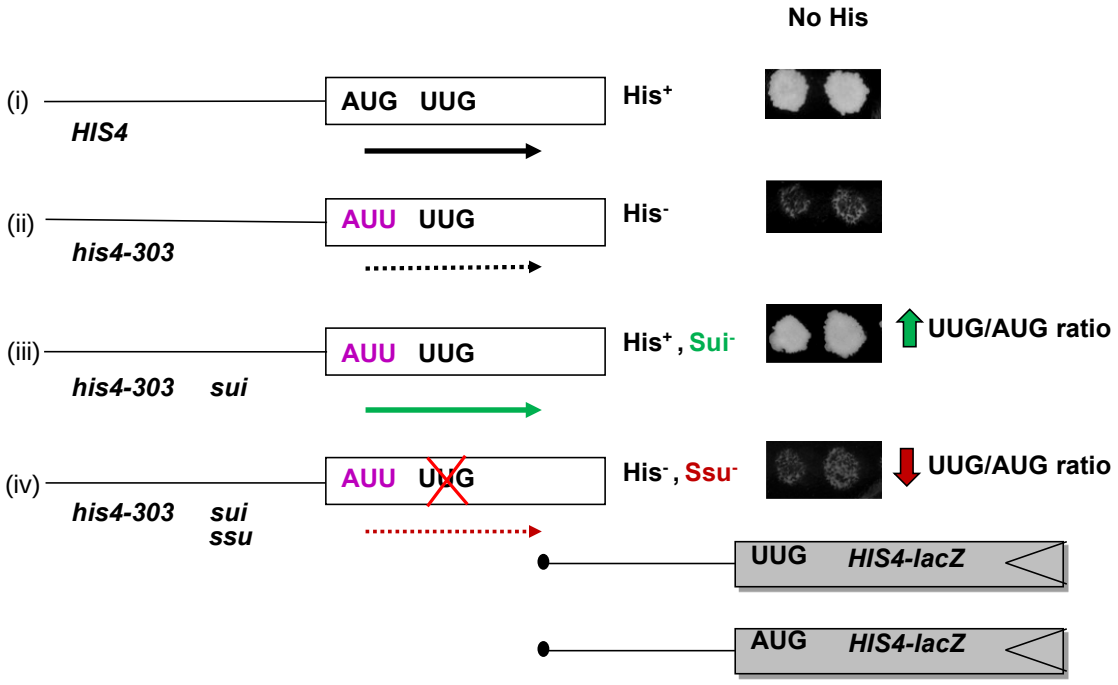

No His
(i)
HIS4
AUG UUG
His+
(ii)
his4-303
AUU UUG
His-
(iii)
his4-303 sui
AUU UUG
His+ , Sui-
UUG/AUG ratio
(iv)
his4-303 sui ssu
AUU UUG
His- , Ssu-
UUG/AUG ratio
UUG HIS4-lacZ
AUG HIS4-lacZ

**Figure 4.6.7. Strains containing *his4-303* used to assay initiation fidelity.**

Depiction of the WT *HIS4* gene and *his4-303*, in which the AUG start codon has been mutated to AUU. (i) Strains containing WT *HIS4* can produce histidine and have a His+ phenotype. (ii) Otherwise WT strains containing *his4-303* are not able to produce *HIS4* and are unable to grow on media lacking histidine. (iii) The introduction of a $Sui^-$ mutation in an initiation factor decreases the fidelity of start codon selection, allowing initiation at a downstream UUG, producing a $His^+$ phenotype. Growth is seen on media lacking histidine, and when assaying paired *HIS4-lacZ* reporters (below), the UUG:AUG ratio is increased. (iv) An $Ssu^-$ mutation suppresses the $Sui^-$ phenotype of a mutation in another factor, abolishing the ability to initiate at UUG downstream and restoring the $His^-$ phenotype of *his4-303*. There is a suppression of growth on –His media, and a decrease in the UUG:AUG ratio.

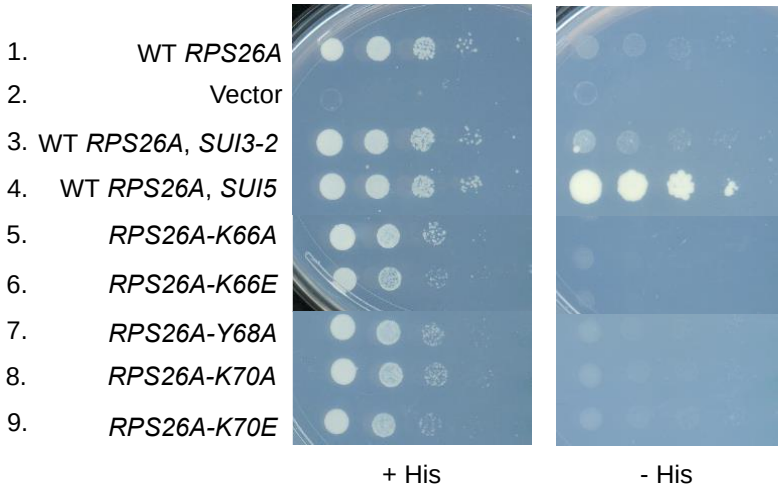



**Figure 4.6.8. Rps26/eS26 variants that contact the mRNA at the -10 nucleotide and further upstream cause growth defects.**

Ten-fold serial dilutions of $P_{GAL1}$-*RPS26 his4-303* strain (JVY09) transformed with the indicated plasmid-born wild type (WT) or mutant *RPS26* alleles, empty vector (B701), SUI3-2 (B4280), or SUI5 (B4281) were spotted on SD + His +Ura (+His) or SD + Ura + 0.1% His (-His) and incubated at 30º C for 3 days and 5 days, respectively.

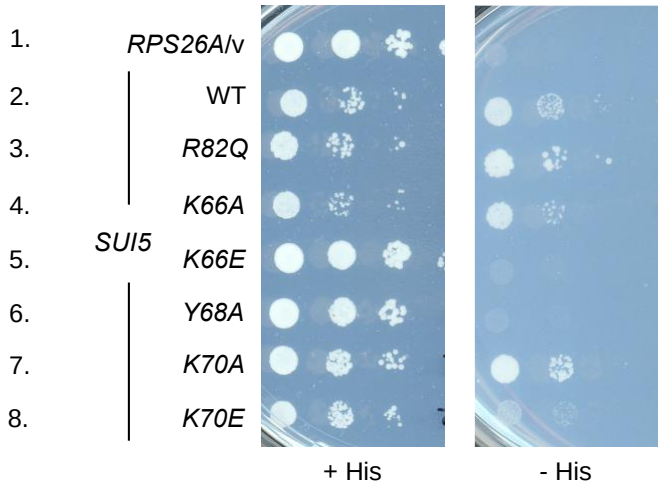



**Figure 4.6.9. Rps26/eS26 variants at positions 66, 68, and 70 confer Ssu$^-$ phenotypes.**

Ten-fold serial dilutions of derivatives of JVY09 containing either empty Trp vector (B701) or *SUI5* (B4281) and the indicated *RPS26A* alleles on high copy *LEU2* plasmids were spotted on SD + His +Ura (+ His) or SD + Ura + 0.1% His (- His) and incubated at 30º C for 3 days and 5 days, respectively.

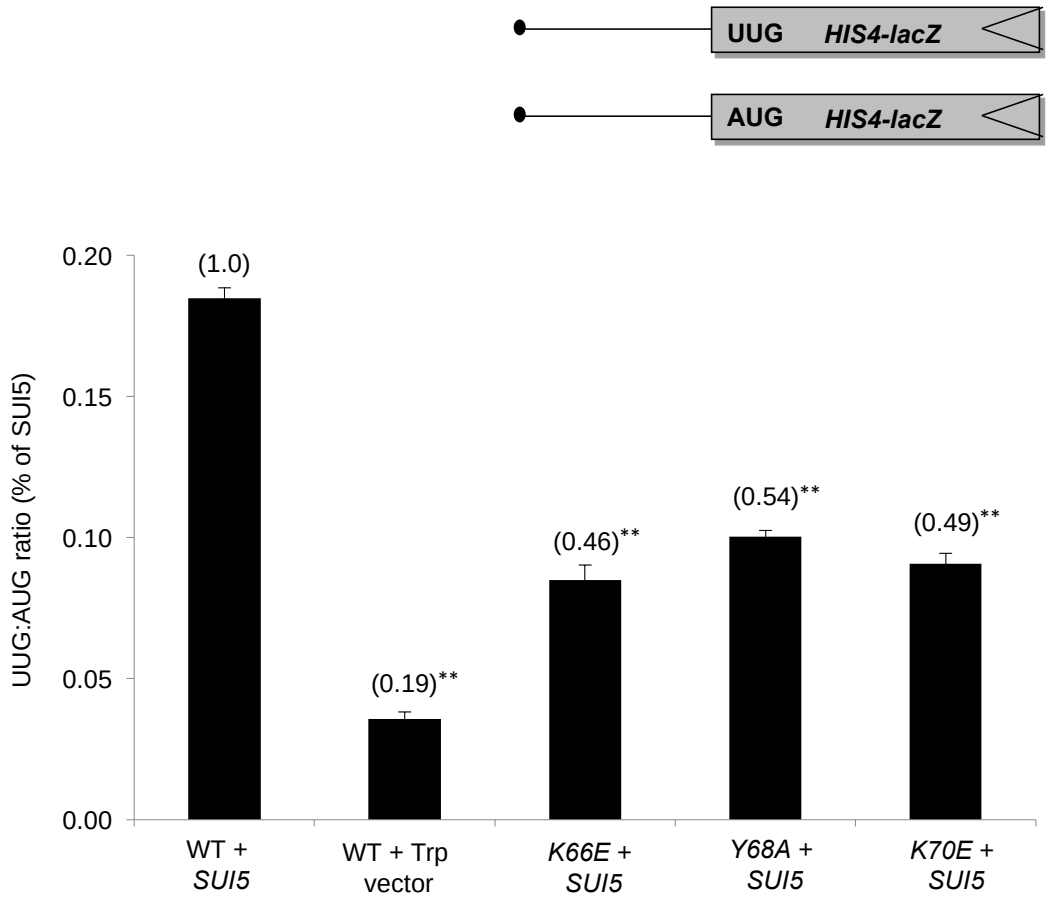

UUG HIS4-lacZ
AUG HIS4-lacZ
UUG:AUG ratio (% of SUI5)
0.20
0.15
0.10
0.05
0.00
(1.0)
(0.19)**
(0.46)**
(0.54)**
(0.49)**
WT + SUI5
WT + Trp vector
K66E + SUI5
Y68A + SUI5
K70E + SUI5

**Figure 4.6.10. Substitutions in Rps26/eS26 residues 66, 68, and 70 increase the stringency of initiation fidelity *in vivo*.**

Derivatives of strain JVY09 containing the indicated *RPS26A* alleles, either Trp vector (B701) or *SUI5* (B4281), and *HIS4-lacZ* reporters with either an AUG or UUG start codon (above) on plasmids p367 (B3989) and p391 (B3990), respectively, were cultured in synthetic complete medium lacking leucine, uracil, and tryptophan (SC –L –U –W) at 30°C to $A_{600}$ of ~0.8, and β-galactosidase activities were measured as in Fig. 4.6.10. The ratio of expression of the UUG to AUG reporter was calculated from at least four different measurements, and the mean and S.E.M.s were plotted. Asterisks indicate significant differences between mutant and wild type, as determined by a two-tailed, unpaired Student's t-test (*$p<0.05$, **$p<0.01$).

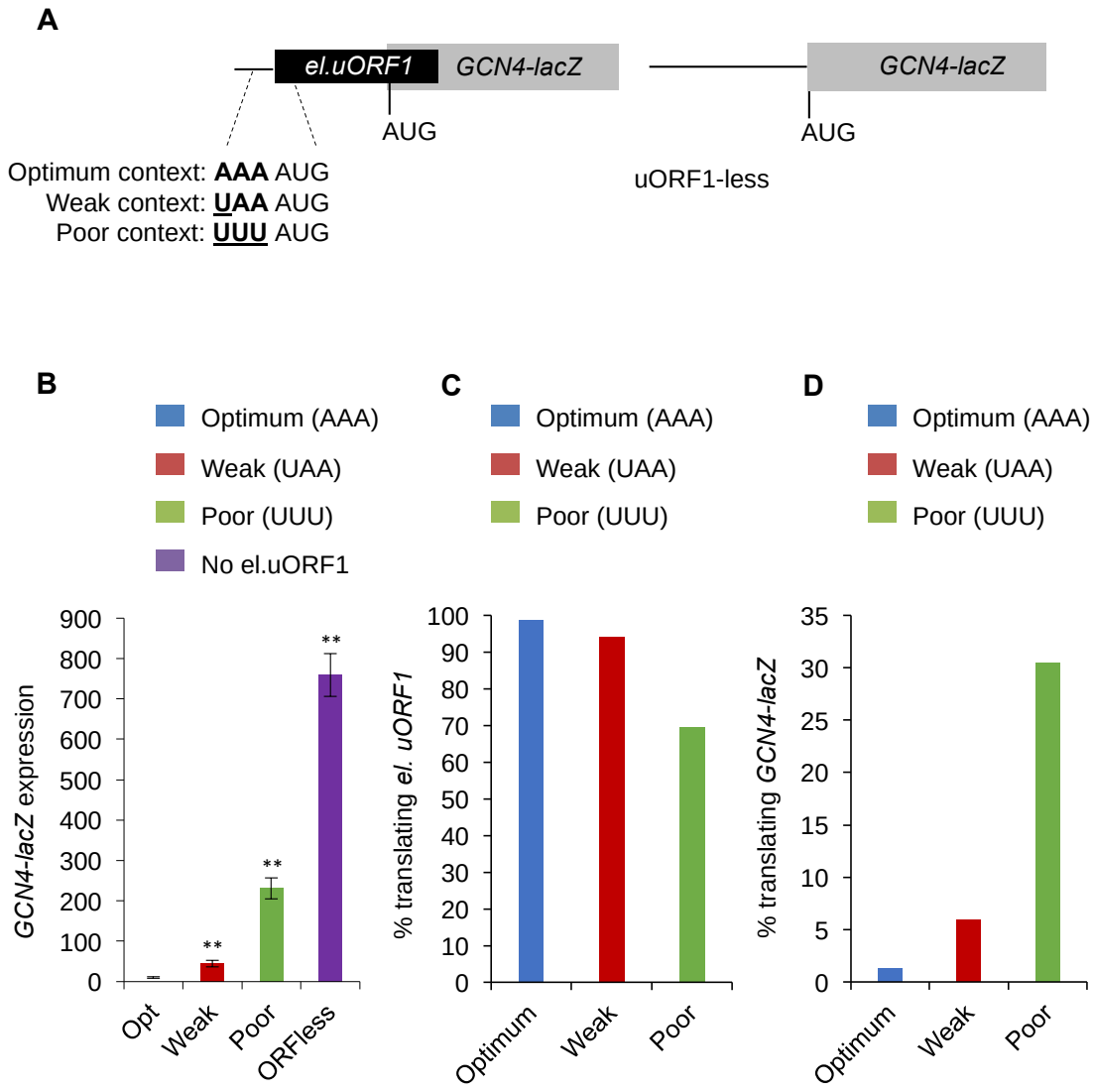
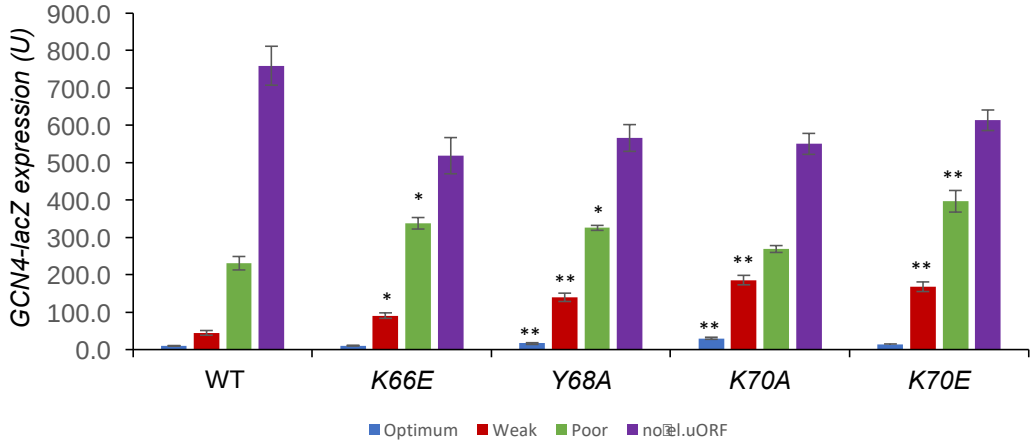

A
el.uORF1
GCN4-lacZ
AUG
GCN4-lacZ
AUG
Optimum context: AAA AUG
Weak context: UAA AUG
Poor context: UUU AUG
uORF1-less
B
Optimum (AAA)
Weak (UAA)
Poor (UUU)
No el.uORF1
GCN4-lacZ expression
900
800
700
600
500
400
300
200
100
0
**
**
**
Opt
Weak
Poor
ORFless
C
Optimum (AAA)
Weak (UAA)
Poor (UUU)
% translating el. uORF1
100
90
80
70
60
50
40
30
20
10
0
Optimum
Weak
Poor
D
Optimum (AAA)
Weak (UAA)
Poor (UUU)
% translating GCN4-lacZ
35
30
25
20
15
10
5
0
Optimum
Weak
Poor

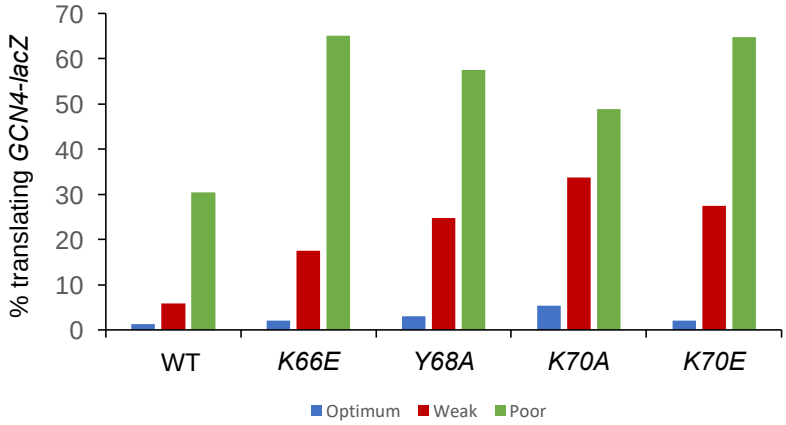
**Figure 4.6.11. Expression and leaky scanning of *GCN4-lacZ* reporters.**

**(A)** Depiction of *GCN4-lacZ* reporters for testing the effects of start codon sequence context on leaky scanning. uORF-1 of the *GCN4* gene has been elongated to overlap the *GCN4* CDS (el.uORF1), with the other uORFs removed, and uAUG-1 is placed in either optimum (AAA), weak (UAA), or poor (UUU) context (left). An ORF-less reporter in which the AUG of uORF1 is mutated is used to estimate maximum initiation (right).
**(B)** Derivatives of strain JVY09 containing the WT *RPS26A* (JVB06) and also harboring the *GCN4-lacZ* reporters depicted in (A) (pC3502, pC3503, pC4466, or pC3505) were cultured in synthetic complete medium lacking leucine and uracil (SC –L –U) at 30°C to $A_{600}$ of ~0.8, and β-galactosidase activities were measured as in Fig. 4.6.10. Mean expression and S.E.M.s were plotted.
**(C, D)** The percentages of scanning ribosomes that translate el.uORF1 (C) or leaky scan uAUG-1 and translate *GCN4-lacZ* (D) were calculated from the data in (B) by normalizing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (D), and subtracting the values in (D) from 100 to obtain the percentages in (C).

**A**


GCN4-lacZ expression (U)
900.0
800.0
700.0
600.0
500.0
400.0
300.0
200.0
100.0
0.0
WT
K66E
Y68A
K70A
K70E
Optimum
Weak
Poor
no el.uORF

**B**


% translating GCN4-lacZ
70
60
50
40
30
20
10
0
WT
K66E
Y68A
K70A
K70E
Optimum
Weak
Poor

**C**

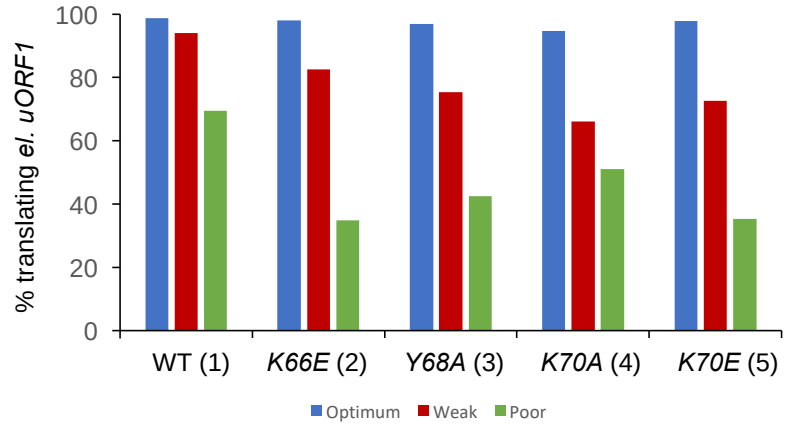

% translating el. uORF1
100
80
60
40
20
0
WT (1)
K66E (2)
Y68A (3)
K70A (4)
K70E (5)
Optimum
Weak
Poor

**Figure 4.6.12. Expression and leaky scanning of *GCN4-lacZ* reporters.**

**(A)** Derivatives of strain JVY09 containing the WT *RPS26A* (JVB06) or the indicated *RPS26A* variants and also harboring the *GCN4-lacZ* reporters depicted in Fig. 4.6.12 A (pC3502, pC3503, pC4466, or pC3505) were cultured in synthetic complete medium lacking leucine and uracil (SC –L –U) at 30°C to $A_{600}$ of ~0.8, and β-galactosidase activities were measured as in Fig. 4.6.10. Mean expression levels and SEMs were plotted.

**(B, C)** The percentages of scanning ribosomes that leaky scan uAUG-1 and translate *GCN4-lacZ* (B) or translate el.uORF1 (C) were calculated from the values in (A) by normalizing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (C), and subtracting the values in (C) from 100 to obtain the percentages in (B). Asterisks indicate significant differences between mutant and wild type, as determined by a two-tailed, unpaired Student's t-test (*$p<0.05$, **$p<0.01$).

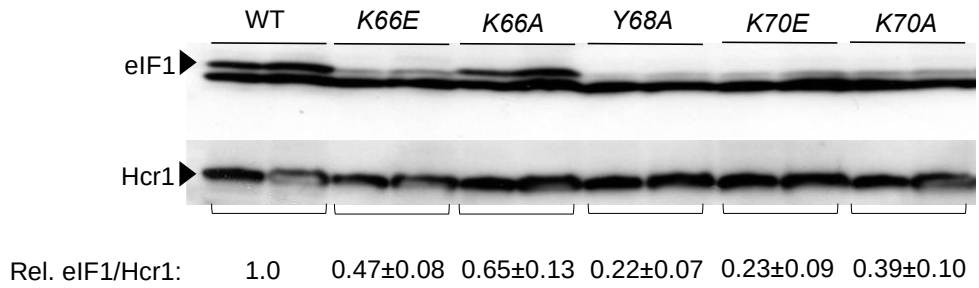
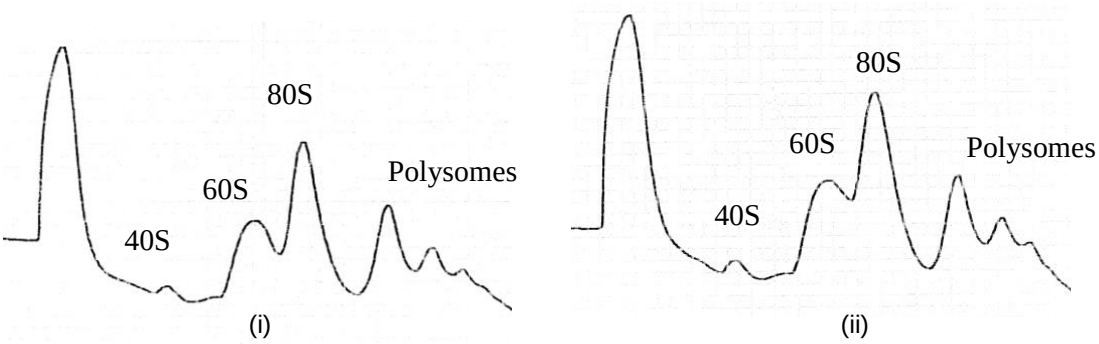



**Figure 4.6.13. Substitutions of Rps26/eS26 residues K66, Y68, and K70 exacerbate poor context at the native *SUI1* AUG to reduce eIF1 expression.**

WCEs of derivatives of $P_{GAL1}$-*RPS26 his4-303* strain (JVY09) transformed with the indicated plasmid-born wild type (WT) or mutant *RPS26* alleles were subjected to Western analysis using antibodies against eIF1 or Hcr1 (as loading control). Two amounts of each extract differing by a factor of two were loaded in successive lanes. Signal intensities were quantified and mean SUI1/Hcr1 ratios with SEMs from three transformants are shown.

**A**




(i) High copy Rps26

40S/60S: 0.13±0.04

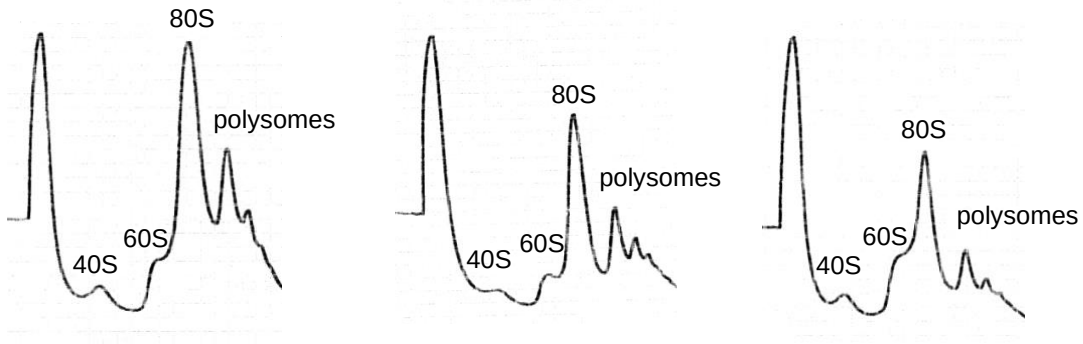
P/M: 1.00±0.12

(ii) High copy *K66E*

40S/60S: 0.16±0.06

P/M: 0.88±0.06

**B**




(i) Parent strain (WT Rps26)

40S/60S: 0.71±0.08

P/M: 0.71±0.12

(ii) Low copy Rps26

40S/60S: 0.68±0.05

P/M: 0.82±0.15

(iii) High copy Rps26

40S/60S: 0.50±0.07

P/M: 0.52±0.09

**Figure 4.6.14. Expression of high copy Rps26/eS26 depresses the 40S:60S ratio.**

**(A)** Derivatives of JVY09 containing WT *RPS26A* or *RPS26A-K66E* on a high copy *LEU2* plasmid (JVB06) were cultured in synthetic glucose medium supplemented with histidine, uracil, and tryptophan (SD + His + Ura + Trp) at 30°C to $A_{600}$ of ~0.8, and cycloheximide was added prior to harvesting. WCEs were separated by sucrose density gradient centrifugation and analyzed at 254 nm. Mean 40S/60S and polysomes:monosomes (P:M) ratios with SEMs from three biological replicates are shown.

**(B)** The parent strain (H2994) containing WT chromosomal copies of both paralogs of Rps26/eS26 (i) and derivatives of JVY09 containing *RPS26A* on either a low copy (JVB05) (ii) or high copy (JVB06) (iii) *LEU2* plasmid were cultured in synthetic glucose medium supplemented with histidine, uracil, and tryptophan (SD + His + Ura + Trp) at 30°C to $A_{600}$ of ~0.8, and cycloheximide was added prior to harvesting. WCEs were separated by sucrose density gradient centrifugation and analyzed at 254 nm. Mean 40S/60S and polysomes:monosomes (P:M) ratios with SEMs from three biological replicates are shown.

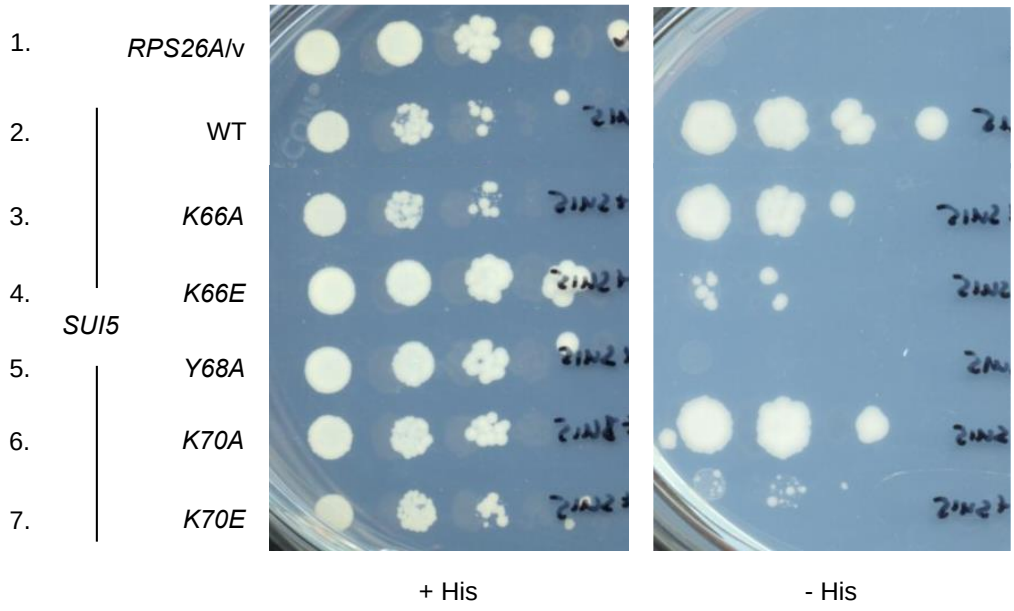
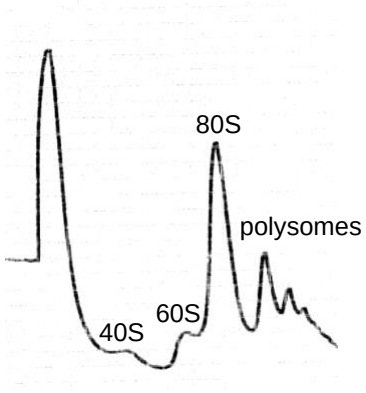



**Figure 4.6.15. Rps26/eS26 variants at positions 66, 68, and 70 on low copy vectors confer Ssu$^-$ phenotypes.**

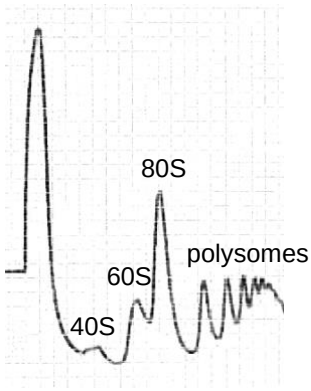
Ten-fold serial dilutions of derivatives of JVY09 containing either empty Trp vector (B701) or *SUI5* (B4281) and the indicated *RPS26A* alleles on low copy *LEU2* plasmids were spotted on SD + His +Ura (+ His) or SD + Ura + 0.1% His (- His) and incubated at 30º C for 3 days and 5 days, respectively.




(i) Low copy WT Rps26

40S/60S: 0.71±0.05

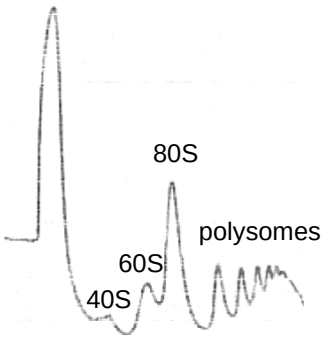
P/M: 0.81±0.10




(ii) *K66E*

40S/60S: 0.33±0.06

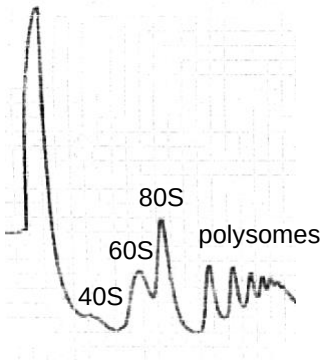
P/M: 1.68±0.14




(iii) *Y68A*

40S/60S: 0.38±0.04

P/M: 1.66±0.09




(iv) *K70E*

40S/60S: 0.33±0.08

P/M: 2.00±0.12

**Figure 4.6.16. Rps26/eS26 mutants lower the 40S:60S ratio.**

Derivatives of JVY09 containing WT *RPS26A* or the indicated variant on a low copy *LEU2* plasmid (JVB05) were cultured in synthetic glucose medium supplemented with histidine, uracil, and tryptophan (SD + His + Ura + Trp) at 30°C to $A_{600}$ of ~0.8, and cycloheximide was added prior to harvesting. WCEs were separated by sucrose density gradient centrifugation and analyzed at 254 nm. Mean 40S/60S and polysomes:monosomes (P:M) ratios with SEMs from three biological replicates are shown.

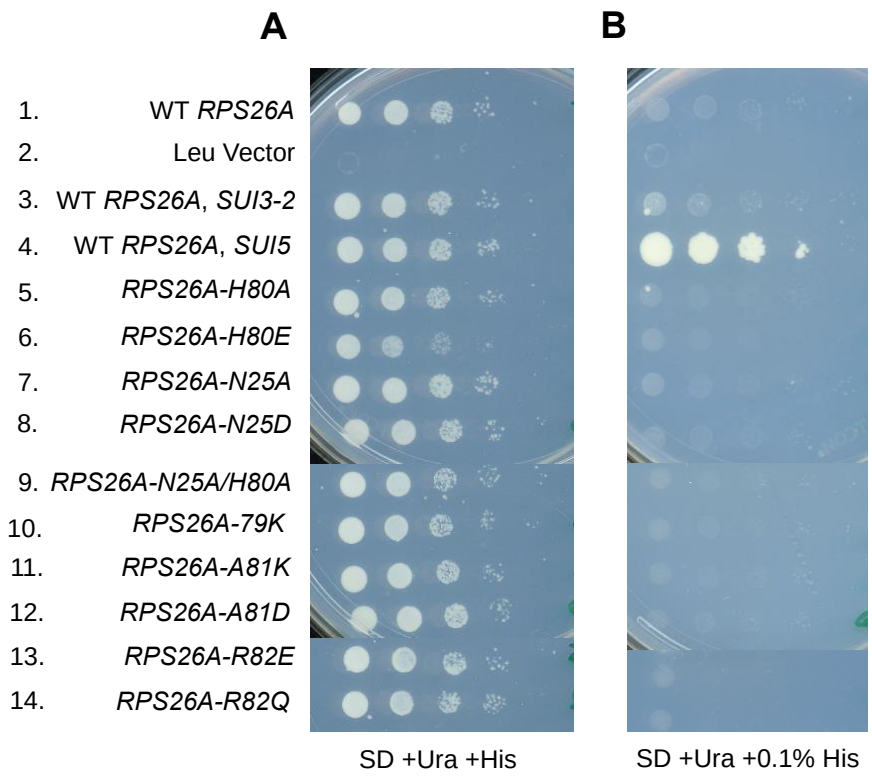
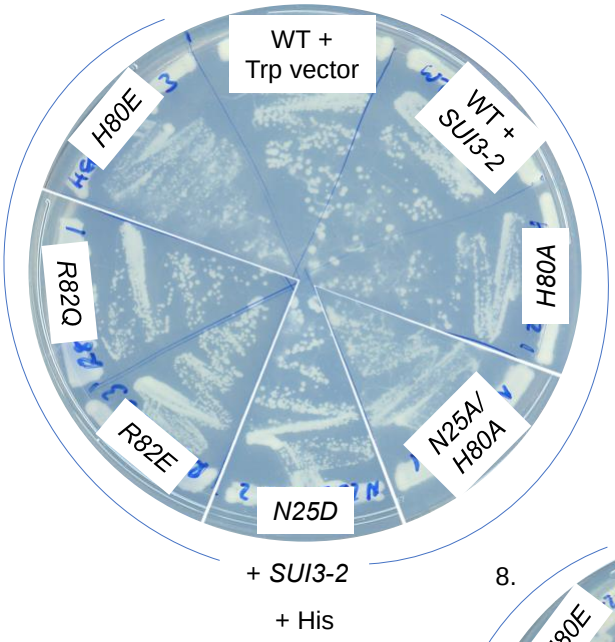



**Figure 4.6.17. Growth of *RPS26* strains with substitutions in α-helix region.**

Ten-fold serial dilutions of $P_{GAL1}$-*RPS26 his4-303* strain (JVY09) transformed with the indicated plasmid-born wild type (WT) or mutant *RPS26* alleles, empty vector (B702), *SUI3-2* (B4280), or *SUI5* (B4281) were spotted on glucose-containing medium, SD + His +Ura (A) or SD + Ura + 0.1% His (B) and incubated at 30º C for 3 days and 5 days, respectively.


WT +
Trp vector
WT +
SUI3-2
H80A
N25A/
H80A
N25D
R82E
R82Q
H80E
+ SUI3-2
+ His

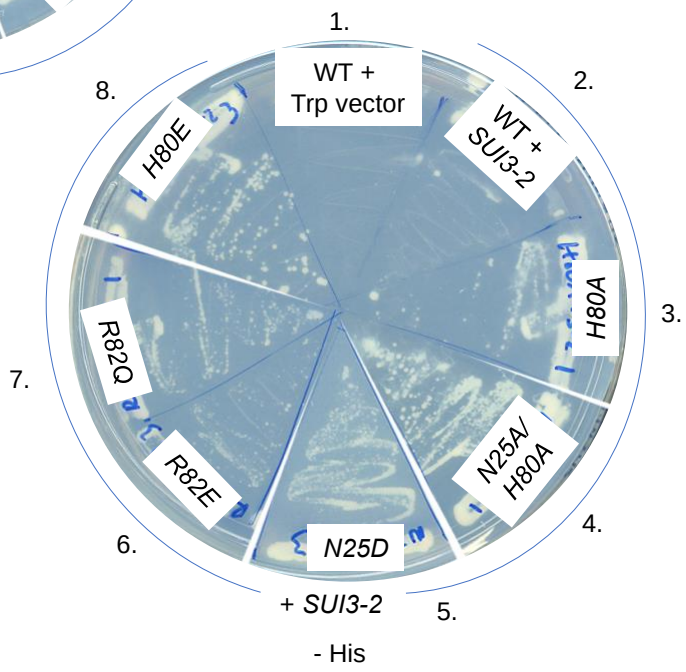

1.
2.
3.
4.
5.
6.
7.
8.
WT +
Trp vector
WT +
SUI3-2
H80A
N25A/
H80A
N25D
R82E
R82Q
H80E
+ SUI3-2
- His

**Figure 4.6.18. Substitutions of Rps26/eS26 residues contacting important context nucleotides confer synthetic His$^+$ phenotypes with *SUI3-2*.**

JVY09 containing either an empty Trp vector (B701) or *SUI3-2* (B4280) and derivatives of JVY09 containing *SUI3-2* and the indicated Rps26 alleles were streaked for growth on SC –L –W (+ His) and SC –L –W –H (– His) and grown at 30º C for 2 days. WT cells are unable to grow on medium lacking histidine because they contain the *his4-303* allele in place of WT *HIS4*. In this strain background, *SUI3-2* does not confer a His$^+$ phenotype. Several Rps26/eS26 variants, which did not facilitate growth on –His medium alone, confer a His$^+$ phenotype in the presence of *SUI3-2*.

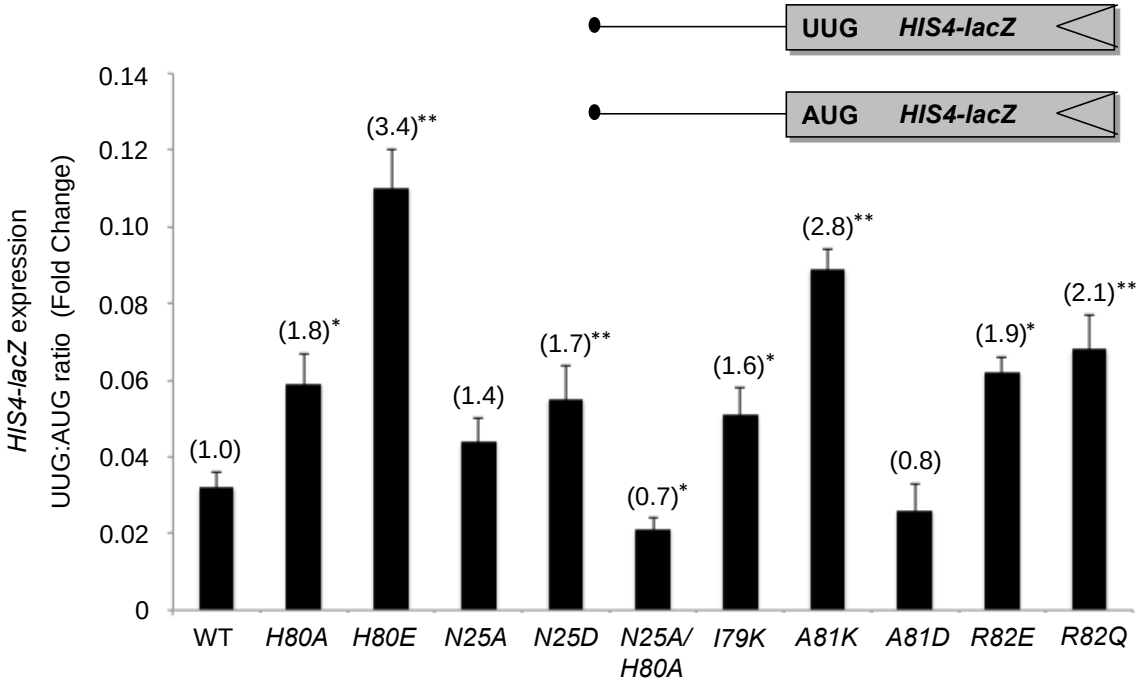

UUG HIS4-lacZ
AUG HIS4-lacZ
HIS4-lacZ expression
UUG:AUG ratio (Fold Change)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
(1.0)
(1.8)*
(3.4)**
(1.4)
(1.7)**
(0.7)*
(1.6)*
(2.8)**
(0.8)
(1.9)*
(2.1)**
WT
H80A
H80E
N25A
N25D
N25A/ H80A
I79K
A81K
A81D
R82E
R82Q

**Figure 4.6.19. Substitutions in Rps26/eS26 residues contacting important context nucleotides decrease initiation fidelity *in vivo*.**

Derivatives of strain JVY09 containing the indicated *RPS26A* alleles and also harboring *HIS4-lacZ* reporters with either an AUG or UUG start codon (above) on plasmids p367 (B3989) and p391 (B3990), respectively, were cultured in synthetic complete medium lacking leucine and uracil (SC –L –U) at 30°C to $A_{600}$ of ~0.8, and β-galactosidase activities (in units of nanomoles of o-nitrophenyl-b-D-galactopyranoside cleaved per min per mg) were measured in whole cell extracts (WCEs). The ratio of expression of the UUG to AUG reporter was calculated from at least four different measurements, and the mean and S.E.M.s were plotted. *NH*, *N25A/H80A*. Asterisks indicate significant differences between mutant and wild type, as determined by a two-tailed, unpaired Student's t-test (*$p<0.05$, **$p<0.01$).

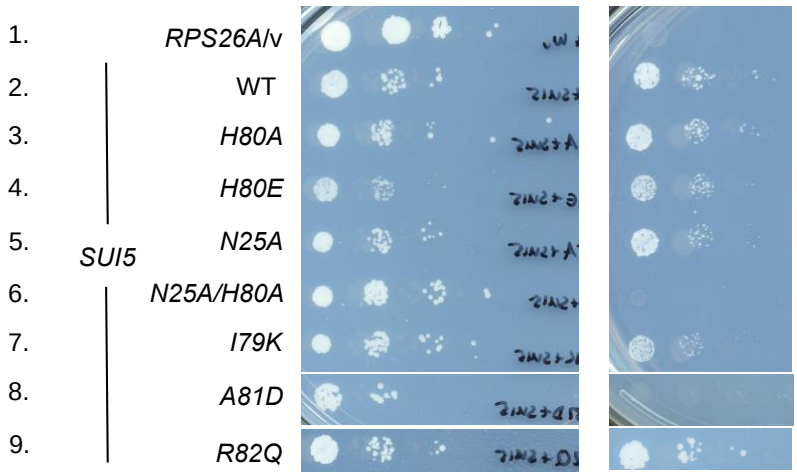
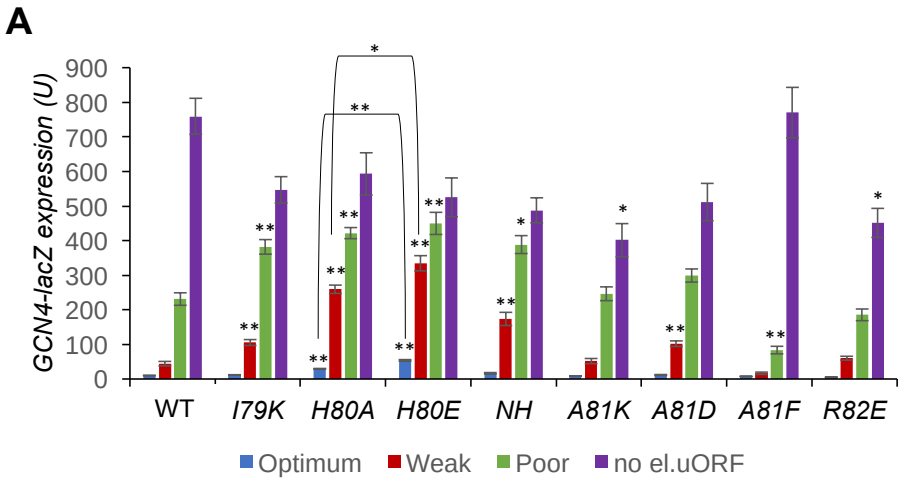
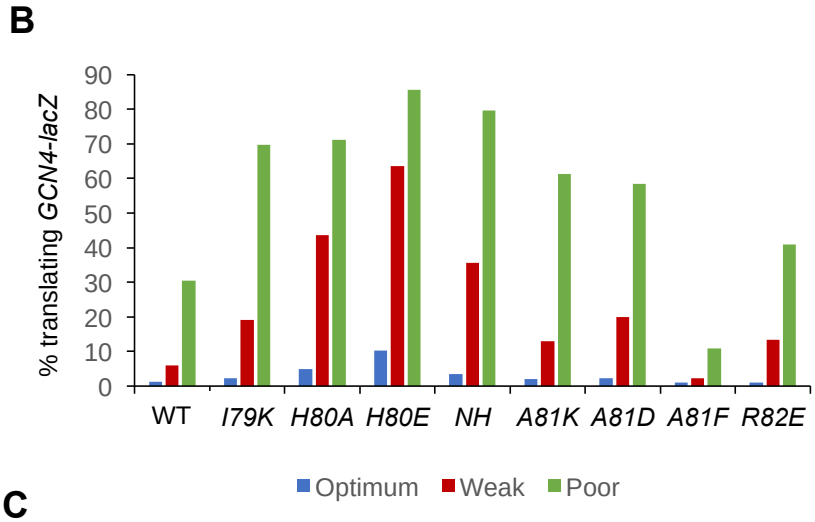

1. RPS26A/v
2. WT
3. H80A
4. H80E
5. N25A
6. N25A/H80A
7. I79K
8. A81D
9. R82Q
SUI5
+ His
0.1% His

**Figure 4.6.20. Rps26/eS26 substitutions N25A/H80A and A81D suppress the His$^+$ phenotype of *SUI5*.**

Ten-fold serial dilutions of derivatives of JVY09 containing either empty Trp vector (B701) or *SUI5* (B4281) and the indicated *RPS26A* alleles on high copy *LEU2* plasmids were spotted on SD + His +Ura (+ His) or SD + Ura + 0.1% His (0.1% His) and incubated at 30º C for 3 days and 5 days, respectively.


A
GCN4-lacZ expression (U)
900
800
700
600
500
400
300
200
100
0
WT
I79K
H80A
H80E
NH
A81K
A81D
A81F
R82E
Optimum
Weak
Poor
no el.uORF


B
% translating GCN4-lacZ
90
80
70
60
50
40
30
20
10
0
WT
I79K
H80A
H80E
NH
A81K
A81D
A81F
R82E
Optimum
Weak
Poor

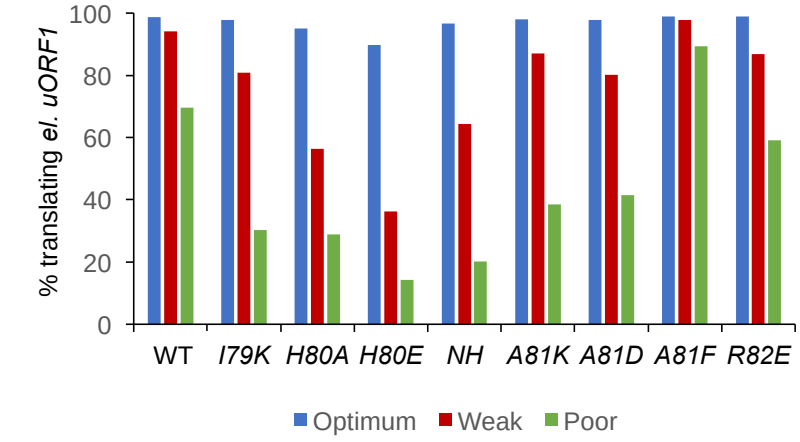

C
% translating el. uORF1
100
80
60
40
20
0
WT
I79K
H80A
H80E
NH
A81K
A81D
A81F
R82E
Optimum
Weak
Poor

**Figure 4.6.21. Substitutions in the α-helix of Rps26/eS26 alter recognition of start codons in suboptimal sequence context.**

**(A)** Derivatives of strain JVY09 containing the WT *RPS26A* (JVB06) or *RPS26A* variants and also harboring the *GCN4-lacZ* reporters depicted in Fig. 4.6.12 A (pC3502, pC3503, pC4466, or pC3505) were cultured in synthetic complete medium lacking leucine and uracil (SC –L –U) at 30°C to $A_{600}$ of ~0.8, and β-galactosidase activities were measured as in Fig. 4.6.10. Mean expression and S.E.M.s were plotted.

**(B, C)** The percentages of scanning ribosomes that leaky scan uAUG-1 and translate *GCN4-lacZ* (B) or that initiate at el.uORF1 (C) were calculated from the data in (A) by normalizing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (B), and subtracting the values in (B) from 100 to obtain the percentages in (C). Asterisks indicate significant differences between mutant and wild type, as determined by a two-tailed, unpaired Student's t-test (*$p<0.05$, **$p<0.01$).

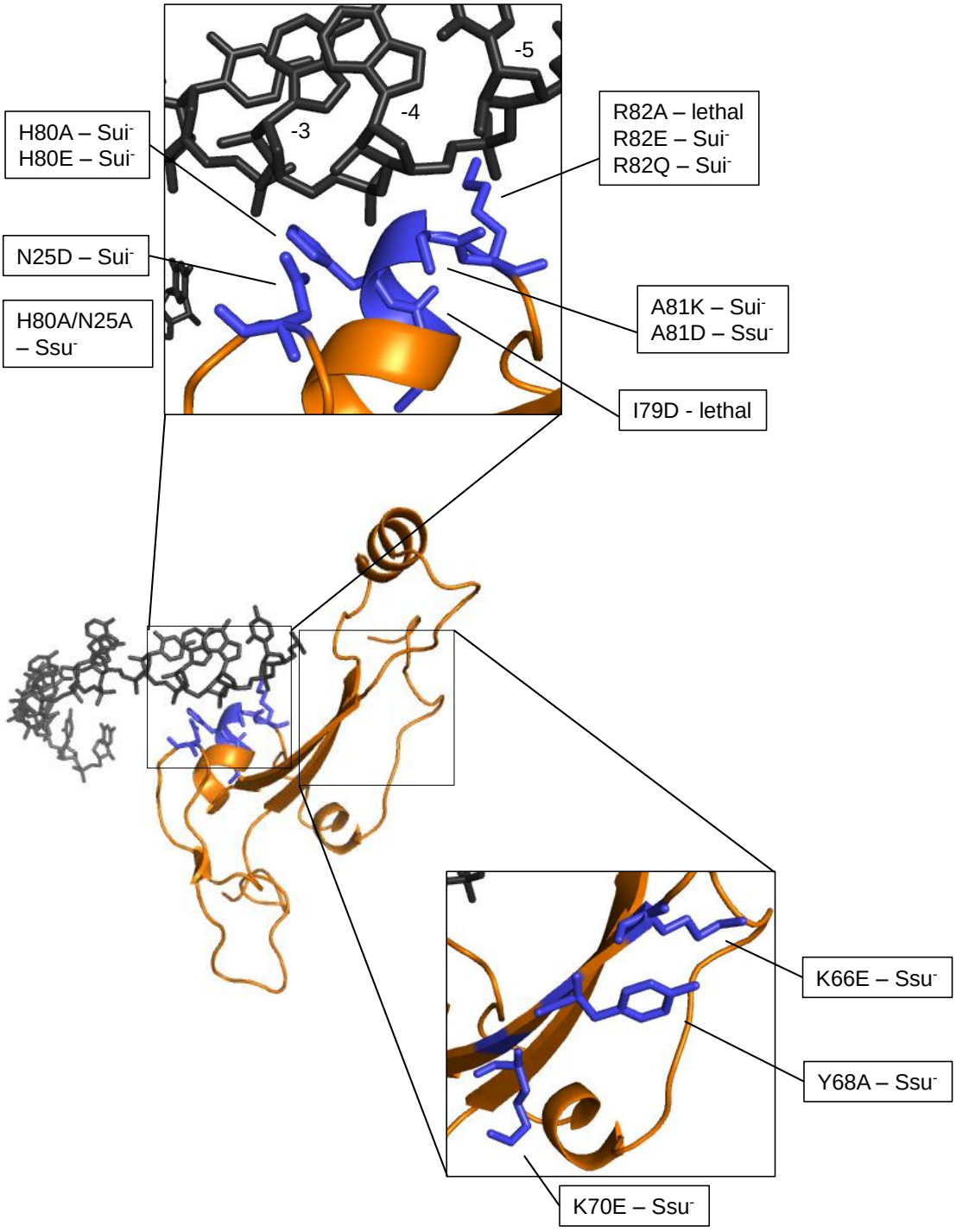

-5
-4
-3
H80A – Sui-
H80E – Sui-
R82A – lethal
R82E – Sui-
R82Q – Sui-
N25D – Sui-
H80A/N25A – Ssu-
A81K – Sui-
A81D – Ssu-
I79D - lethal
K66E – Ssu-
Y68A – Ssu-
K70E – Ssu-

**Figure 4.6.22. Effects on start codon selection by Rps26/eS26 variants.**

Depiction of partial 48S yeast ribosome (PDB 3JAP) showing Rps26 in orange and mRNA in black. Residues mutated in this study are shown as sticks and colored blue, with the phenotypes conferred by each indicated.

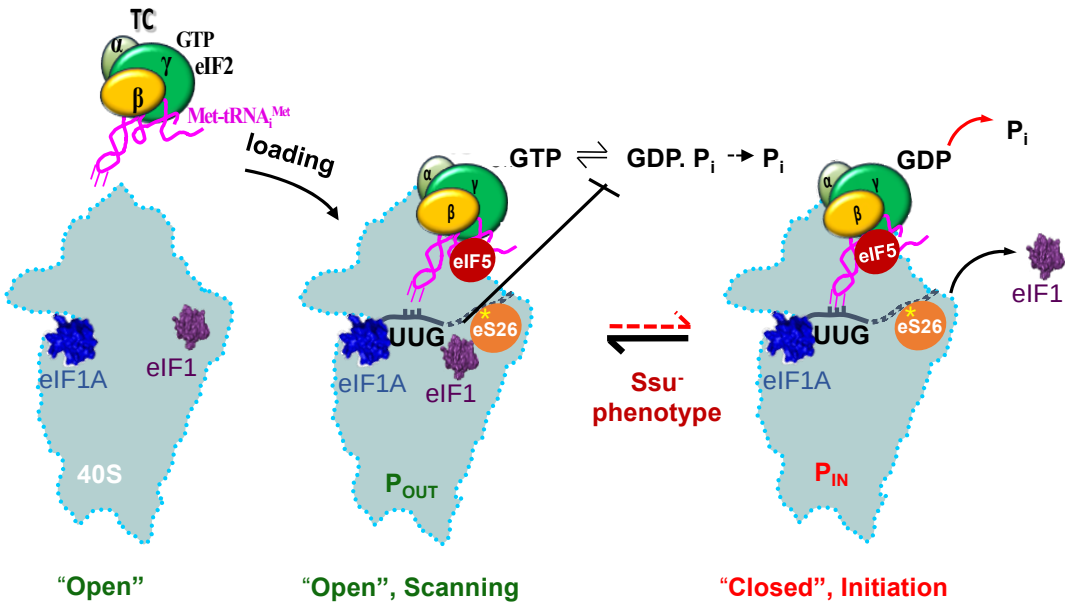



**Figure 4.6.23. Model describing the effects of Rps26/eS26 interactions with the -10 nucleotide and upstream nucleotides in the mRNA exit channel on the conformational rearrangements of the PIC during start codon selection.**

eIF1 and eIF1A promote an open, scanning conformation of the PIC to which TC loads (i). Following TC binding, the PIC scans the mRNA leader in the open conformation (ii). When a start codon is recognized, the PIC transitions to the closed state, accompanied by release of eIF1 and $P_i$ (iii). Mutations that destabilize the closed complex, such as those that remove contacts between eS26 and mRNA, shift the equilibrium of ribosomal complexes toward the open state, increasing the stringency of the requirement for an AUG in good context and decreasing initiation at near-cognate UUG codons. Such mutations confer an Ssu$^-$ phenotype.

**Table 4.6.1**. **Yeast strains used in this study.**

| Strain | Genotype | Source |
|---|---|---|
| HLV01a (H2994) | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU)* | (146) |
| JVY05 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) rps26bΔ::hphMX6* | Jyothsna Visweswaraiah |
| JVY09 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* | Jyothsna Visweswaraiah |
| LMY84 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP5*(*hc *RPS26A-N25A, LEU2)* | This study |
| LMY85 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP6*(*hc *RPS26A-N25K, LEU2)* | This study |
| LMY86 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ:hphMX6* pLMP7*(*hc *RPS26A-N25D, LEU2)* | This study |
| LMY87 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP8*(*hc *RPS26A-K66A, LEU2)* | This study |
| LMY88 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP9*(*hc *RPS26A-K66E, LEU2)* | This study |
| LMY89 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP10*(*hc *RPS26A-Y68A, LEU2)* | This study |
| LMY90 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP11*(*hc *RPS26A-K70A, LEU2)* | This study |
| LMY91 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP12*(*hc *RPS26A-K70E, LEU2)* | This study |
| LMY92 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP13*(*hc *RPS26A-I79K, LEU2)* | This study |
| LMY93 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP14*(*hc *RPS26A-I79D, LEU2)* | This study |
| LMY94 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP15*(*hc *RPS26A-H80A, LEU2)* | This study |

| Strain | Genotype | Source |
|---|---|---|
| LMY95 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ:hphMX6* pLMP16*(*hc *RPS26A-H80E, LEU2)* | This study |
| LMY96 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP17*(*hc *RPS26A-A81F, LEU2)* | This study |
| LMY97 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP18*(*hc *RPS26A-A81K, LEU2)* | This study |
| LMY98 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP19*(*hc *RPS26A-A81D, LEU2)* | This study |
| LMY99 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP20*(*hc *RPS26A-R82A, LEU2)* | This study |
| LMY100 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP21*(*hc *RPS26A-R82Q, LEU2)* | This study |
| LMY101 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP22*(*hc *RPS26A-R82E, LEU2)* | This study |
| LMY102 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP23*(*hc *RPS26A-N25A/H80A, LEU2)* | This study |
| LMY143 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pJVB06*(*hc *RPS26A, LEU2)* | This study |
| LMY144 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pJVB05*(*lc *RPS26A, LEU2)* | This study |
| LMY145 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP75*(*lc *RPS26A-K66A, LEU2)* | This study |
| LMY146 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP76*(*lc *RPS26A-K66E, LEU2)* | This study |
| LMY147 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP77*(*lc *RPS26A-Y68A, LEU2)* | This study |
| LMY148 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::P$_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP78*(*lc *RPS26A-K70A, LEU2)* | This study |

| Strain | Genotype | Source |
|---|---|---|
| LMY149 | *MATa ura3-52 leu2-3 leu2-112 trp1-Δ63 his4-303(AUU) kanMX6::$P_{GAL1}$-RPS26A rps26bΔ::hphMX6* pLMP79(lc *RPS26A-K70E, LEU2)* | This study |

**Table 4.6.2. Plasmids used in this study.**

| Plasmid | Description | Parent Plasmid | Source or Reference |
|---|---|---|---|
| pRS425 (B1378) | hc *LEU2* | | (359) |
| pJVB06 | hc *RPS26A, LEU2* | pRS425 | Jyothsna Visweswaraiah |
| pLMP5 | hc *RPS26A-N25A, LEU2* | JVB06 | This study |
| pLMP6 | hc *RPS26A-N25K, LEU2* | JVB06 | This study |
| pLMP7 | hc *RPS26A-N25D, LEU2* | JVB06 | This study |
| pLMP8 | hc *RPS26A-K66A, LEU2* | JVB06 | This study |
| pLMP9 | hc *RPS26A-K66E, LEU2* | JVB06 | This study |
| pLMP10 | hc *RPS26A-Y68A, LEU2* | JVB06 | This study |
| pLMP11 | hc *RPS26A-K70A, LEU2* | JVB06 | This study |
| pLMP12 | hc *RPS26A-K70E, LEU2* | JVB06 | This study |
| pLMP13 | hc *RPS26A-I79K, LEU2* | JVB06 | This study |
| pLMP14 | hc *RPS26A-I79D, LEU2* | JVB06 | This study |
| pLMP15 | hc *RPS26A-H80A, LEU2* | JVB06 | This study |
| pLMP16 | hc *RPS26A-H80E, LEU2* | JVB06 | This study |
| pLMP17 | hc *RPS26A-A81F, LEU2* | JVB06 | This study |
| pLMP18 | hc *RPS26A-A81K, LEU2* | JVB06 | This study |
| pLMP19 | hc *RPS26A-A81D, LEU2* | JVB06 | This study |
| pLMP20 | hc *RPS26A-R82A, LEU2* | JVB06 | This study |
| pLMP21 | hc *RPS26A-R82Q, LEU2* | JVB06 | This study |
| pLMP22 | hc *RPS26A-R82E, LEU2* | JVB06 | This study |
| pLMP23 | hc *RPS26A-N25A/H80A, LEU2* | JVB06 | This study |
| pRS315 (B702) | lc *LEU2* | pRS305 | (355) |
| pJVB05 | lc *RPS26A, LEU2* | pRS315 | Jyothsna Visweswaraiah |
| pLMP75 | lc *RPS26A-K66A, LEU2* | JVB05 | This study |

| Plasmid | Description | Parent Plasmid | Source or Reference |
|---|---|---|---|
| pLMP76 | lc *RPS26A-K66E, LEU2* | JVB05 | This study |
| pLMP77 | lc *RPS26A-Y68A, LEU2* | JVB05 | This study |
| pLMP78 | lc *RPS26A-K70A, LEU2* | JVB05 | This study |
| pLMP79 | lc *RPS26A-K70E, LEU2* | JVB05 | This study |
| pRS314 (B701) | lc *TRP1* | pRS304 | (355) |
| YCpSUI3-S264Y-W (B4280) | sc *SUI3-S264Y, TRP1* | YCplac22 | (146) |
| YCpTIF5-G31R-W (B4281) | sc *TIF5-G31R, TRP1* | YCplac22 | (146) |
| p367 (B3989) | sc *URA3 HIS4(ATG)-lacZ* | p349 | (360) |
| p391 (B3990) | sc *URA3 HIS4(TTG)-lacZ* | p349 | (360) |
| pC3502 | sc *URA3* $^{-3}$AAA$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |
| pC3503 | sc *URA3* $^{-3}$UUU$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |
| pC3505 | sc *URA3* el.uORF1-less *GCN4-lacZ* in YCp50 | | (314) |
| pC4466 | sc *URA3* $^{-3}$UAA$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |

**Table 4.6.3. Oligonucleotide primers used in this study.**

| SUBSTITUTION | SEQUENCE 5' -3' |
|---|---|
| **RPS26A-N25A-FOR** | 5'-TTTGGAATAGACTTGGAACAGGCGACACATCTGACTGGTTTGAC-3' |
| **RPS26A-N25A-REV** | 5'-GTCAAACCAGTCAGATGTGTCGCCTGTTCCAAGTCTATTCCAAA-3' |
| **RPS26A-N25K-FOR** | 5'-AAACCAGTCAGATGTGTCAAGTGTTCCAAGTCTATTCC-3' |
| **RPS26A-N25K-REV** | 5'-GGAATAGACTTGGAACACTTGACACATCTGACTGGTTT-3' |
| **RPS26A-N25D-FOR** | 5'-GGAATAGACTTGGAACAGTCGACACATCTGACTGGTTTG-3' |
| **RPS26A-N25D-REV** | 5'-CAAACCAGTCAGATGTGTCGACTGTTCCAAGTCTATTCC-3' |
| **RPS26A-K66A-FOR** | 5'-GTAGTGTAACTTGTTGTAAGTCGCTGGCAAAGCGTATTCAGGGTAG-3' |
| **RPS26A-K66A-REV** | 5'-CTACCCTGAATACGCTTTGCCAGCGACTTACAACAAGTTACACTAC-3' |
| **RPS26A-K66E-FOR** | 5'-GTAACTTGTTGTAAGTCTCTGGCAAAGCGTATTCAGGGTA-3' |
| **RPS26A-K66E-REV** | 5'-TACCCTGAATACGCTTTGCCAGAGACTTACAACAAGTTAC-3' |
| **RPS26A-Y68A-FOR** | 5'-AAACACAGTAGTGTAACTTGTTGGCAGTCTTTGGCAAAGCGTATTCAG-3' |
| **RPS26A-Y68A-REV** | 5'-CTGAATACGCTTTGCCAAAGACTGCCAACAAGTTACACTACTGTGTTT-3' |
| **RPS26A-K70A-FOR** | 5'-ACAAGAAACACAGTAGTGTAACGCGTTGTAAGTCTTTGGCAAAGCG-3' |

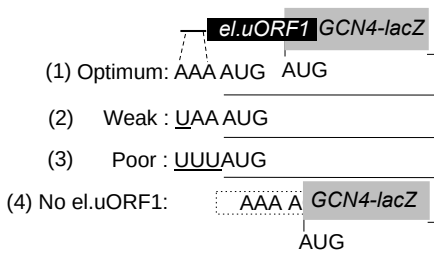
| SUBSTITUTION | SEQUENCE 5’ -3’ |
|---|---|
| **RPS26A-K70A-REV** | 5'-CGCTTTGCCAAAGACTTACAACGCGTTACACTACTGTGTTTCTTGT-3' |
| **RPS26A-K70E-FOR** | 5'-AAGAAACACAGTAGTGTAACTCGTTGTAAGTCTTTGGCAAAGC-3' |
| **RPS26A-K70E-REV** | 5'-GCTTTGCCAAAGACTTACAACGAGTTACACTACTGTGTTTCTT-3' |
| **RPS26A-I79K-FOR** | 5'-GACAATTCTGGCGTGCTTAGCACAAGAAACACAGTAGTGTAACTT-3' |
| **RPS26A-I79K-REV** | 5'-AAGTTACACTACTGTGTTTCTTGTGCTAAGCACGCCAGAATTGTC-3' |
| **RPS26A-I79D-FOR** | 5'-GACAATTCTGGCGTGATCAGCACAAGAAACACAGTAGTGTAAC-3' |
| **RPS26A-I79D-REV** | 5'-GTTACACTACTGTGTTTCTTGTGCTGATCACGCCAGAATTGTC-3' |
| **RPS26A-H80A-FOR** | 5'-CTGACAATTCTGGCGGCAATAGCACAAGAAACACAGTAGTG-3' |
| **RPS26A-H80A-REV** | 5'-CACTACTGTGTTTCTTGTGCTATTGCCGCCAGAATTGTCAG-3' |
| **RPS26A-H80E-FOR** | 5'-CTCTGACAATTCTGGCCTCAATAGCACAAGAAACACAGTAGTG-3' |
| **RPS26A-H80E-REV** | 5'-CACTACTGTGTTTCTTGTGCTATTGAGGCCAGAATTGTCAGAG-3' |
| **RPS26A-A81F-FOR** | 5'-GGATCTGACTCTGACAATTCTGAAGTGAATAGCACAAGAAACACAG-3' |
| **RPS26A-A81F-REV** | 5'-CTGTGTTTCTTGTGCTATTCACTTCAGAATTGTCAGAGTCAGATCC-3' |
| **RPS26A-A81K-FOR** | 5'-CTGGATCTGACTCTGACAATTCTCTTGTGAATAGCACAAGAAACACAGT-3' |

| SUBSTITUTION | SEQUENCE 5’ -3’ |
|---|---|
| **RPS26A-A81D-FOR** | 5'-TCTGACTCTGACAATTCTGTCGTGAATAGCACAAGAAAC-3' |
| **RPS26A-A81D-REV** | 5'-GTTTCTTGTGCTATTCACGACAGAATTGTCAGAGTCAGA-3' |
| **RPS26A-R82A-FOR** | 5'-GGATCTGACTCTGACAATTGCGGCGTGAATAGCACAAGAA-3' |
| **RPS26A-R82A-REV** | 5'-TTCTTGTGCTATTCACGCCGCAATTGTCAGAGTCAGATCC-3' |
| **RPS26A-R82Q-FOR** | 5'-CTCTGGATCTGACTCTGACAATCTGGGCGTGAATAGCACAAGAAACA-3' |
| **RPS26A-R82Q-REV** | 5'-TGTTTCTTGTGCTATTCACGCCCAGATTGTCAGAGTCAGATCCAGAG-3' |
| **RPS26A-R82E-FOR** | 5'-CTCTGGATCTGACTCTGACAATCTCGGCGTGAATAGCACAAGAAACA-3' |
| **RPS26A-R82E-REV** | 5'-TGTTTCTTGTGCTATTCACGCCGAGATTGTCAGAGTCAGATCCAGAG-3' |

*FOR, forward primer ; REV, reverse primer

**A**

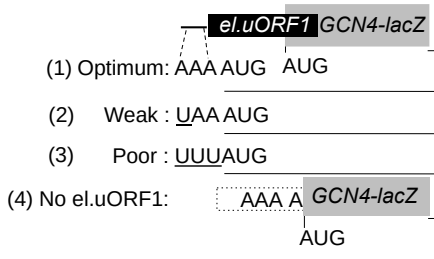
*GCN4-lacZ* expression(U)




| | (1) WT | (2) *K66E* | (3) *Y68A* | (4) K70A | (5) *K70E* |
|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG | 10±0.8 | 10±0.9 | 17±1.0 | 29±3.8 | 13±2.1 |
| (2) Weak : UAA AUG | 45±7 | 91±7 | 140±11 | 186±13 | 168±13 |
| (3) Poor : UUUAUG | 231±18 | 300±39 | 307±17 | 257±10 | 367±42 |
| (4) No el.uORF1: AAA A GCN4-lacZ AUG | 760±52 | 519±48 | 566±36 | 550±28 | 614±28 |

**B**

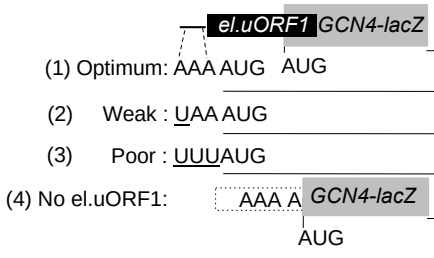
% translating *GCN4-lacZ*




| | (1) WT | (2) *K66E* | (3) *Y68A* | (4) K70A | (5) *K70E* |
|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG | 1.3 | 2.0 | 3.1 | 5.3 | 2.1 |
| (2) Weak : UAA AUG | 5.9 | 17.5 | 24.7 | 33.8 | 27.4 |
| (3) Poor : UUUAUG | 30.4 | 57.8 | 54.2 | 46.7 | 59.7 |
| (4) No el.uORF1: AAA A GCN4-lacZ AUG | >99 | >99 | >99 | >99 | >99 |

**C**

% translating *el. uORF1*




| | (1) WT | (2) *K66E* | (3) *Y68A* | (4) K70A | (5) *K70E* |
|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG | 98.7 | 98.0 | 96.9 | 94.7 | 97.9 |
| (2) Weak : UAA AUG | 94.1 | 82.5 | 75.3 | 66.3 | 72.6 |
| (3) Poor : UUUAUG | 69.6 | 42.3 | 45.8 | 53.3 | 40.3 |
| (4) No el.uORF1: AAA A GCN4-lacZ AUG | >1 | >1 | >1 | >1 | >1 |

**Table 4.6.4. Expression and leaky scanning of *GCN4-lacZ* reporters.**

**(A)** Derivatives of strain JVY09 containing the WT *RPS26A* (JVB06) or *RPS26A* variants and also harboring the *GCN4-lacZ* reporters depicted in Fig. 4.6.12 A (pC3502, pC3503, pC4466, or pC3505) were assayed as described in Fig. 4.6.17 A. Mean expression and S.E.M.s are listed.

**(B, C)** The percentages of scanning ribosomes that leaky scan uAUG-1 and translate *GCN4-lacZ* (B) or that initiate at el.uORF1 (C) were calculated from the data in (A) by normalizing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (B), and subtracting the values in (B) from 100 to obtain the percentages in (C).

**A**

*GCN4-lacZ* expression(U)

| | (1) WT | (2) *I79K* | (3) *H80A* | (4) H80E | (5) *NH* | (6) *A81K* | (7) A81D | (8) *A81F* | (9) *R82E* |
|---|---|---|---|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG *el.uORF1* *GCN4-lacZ* | 10±0.8 | 12±0.6 | 29±1.2 | 54±2.3 | 17±2.1 | 9±0.5 | 11±1.4 | 8±0.3 | 5±0.6 |
| (2) Weak : UAA AUG | 45±7 | 105±8 | 259±12 | 334±22 | 174±19 | 52±7 | 102±8 | 17±3 | 60±6 |
| (3) Poor : UUUAUG | 231±18 | 382±21 | 422±16 | 450±32 | 389±26 | 246±20 | 299±19 | 84±11 | 185±17 |
| (4) No el.uORF1: AAA A *GCN4-lacZ* AUG | 760±52 | 547±38 | 593±61 | 525±56 | 488±36 | 401±48 | 511±54 | 770±73 | 452±42 |

**B**

% translating *GCN4-lacZ*

| | (1) WT | (2) *I79K* | (3) *H80A* | (4) H80E | (5) *NH* | (6) *A81K* | (7) A81D | (8) *A81F* | (9) *R82E* |
|---|---|---|---|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG *el.uORF1* *GCN4-lacZ* | 1.3 | 2.2 | 4.9 | 10.2 | 3.4 | 2.1 | 2.2 | 1.0 | 1.1 |
| (2) Weak : UAA AUG | 5.9 | 19.2 | 43.7 | 63.7 | 35.6 | 12.9 | 19.9 | 2.2 | 13.3 |
| (3) Poor : UUUAUG | 30.4 | 69.8 | 71.1 | 85.7 | 79.7 | 61.4 | 58.5 | 10.8 | 41.0 |
| (4) No el.uORF1: AAA A *GCN4-lacZ* AUG | >99 | >99 | >99 | >99 | >99 | >99 | >99 | >99 | >99 |

**C**

% translating *el. uORF1*

| | (1) WT | (2) *I79K* | (3) *H80A* | (4) H80E | (5) *NH* | (6) *A81K* | (7) A81D | (8) *A81F* | (9) *R82E* |
|---|---|---|---|---|---|---|---|---|---|
| (1) Optimum: AAA AUG AUG *el.uORF1* *GCN4-lacZ* | 98.7 | 97.8 | 95.1 | 89.8 | 96.6 | 97.9 | 97.8 | 99.0 | 98.9 |
| (2) Weak : UAA AUG | 94.1 | 80.8 | 56.3 | 36.3 | 64.4 | 87.1 | 80.1 | 97.8 | 86.7 |
| (3) Poor : UUUAUG | 69.6 | 30.2 | 28.9 | 14.3 | 20.3 | 38.6 | 41.5 | 89.2 | 59.0 |
| (4) No el.uORF1: AAA A *GCN4-lacZ* AUG | <1 | <1 | <1 | <1 | <1 | <1 | <1 | <1 | <1 |

**Table 4.6.5. Expression and leaky scanning of *GCN4-lacZ* reporters.**

**(A)** Derivatives of strain JVY09 containing the WT *RPS26A* (JVB06) or *RPS26A* variants and also harboring the *GCN4-lacZ* reporters depicted in Fig. 4.6.12A (pC3502, pC3503, pC4466, or pC3505) were assayed as described in Fig. 4.6.13A. Mean expression and S.E.M.s are listed.

**(B, C)** The percentages of scanning ribosomes that leaky scan uAUG-1 and translate *GCN4-lacZ* (B) or that initiate at el.uORF1 (C) were calculated from the data in (A) by normalizing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (B), and subtracting the values in (B) from 100 to obtain the percentages in (C).

# 5. Conclusion

Translation initiation is a critical step in protein expression that determines 1) which proteins will be made, 2) which isoforms of those proteins will be expressed, 3) how much of each protein will be synthesized, and also 4) sets the reading frame to ensure synthesis of the correct proteins. Perturbations in this process significantly alter the proteome and cause disease in humans. The scanning eukaryotic ribosome requires complicated molecular machinery to ensure selection of the correct start codon with the correct frequency. Even minor (i.e., single residue/nucleotide) changes in the factors involved can alter the fidelity of initiation enough to cause growth defects in yeast. It is important to note that changes in fidelity in either direction can be detrimental. Start codon selection in eukaryotes has been tuned to allow occasional initiation at non-AUG codons, AUG codons in poor context, and other non-optimal start codons, and these apparent infidelities are not accidental, but rather play important regulatory roles in determining the proteome. There is much evidence that such alternative initiation events allow cells to respond rapidly to stressors or developmental signals. Scanning ribosomes also skip over certain AUG codons, and alterations that prevent such leaky scanning can be detrimental. Thus, we see translation initiation as a dynamic process, carefully calibrated for the stringency of start codon selection, that plays an outsized role in determining the proteome of the cell and rapidly changing that proteome in response to changing conditions.

At the start of this study, it was hypothesized that the 40S ribosome, with the aid of eIF1 and eIF1A, assumed an open conformation that facilitated mRNA recruitment and scanning before transitioning to a closed conformation upon AUG recognition. Structural evidence for this model was mixed, however, creating a controversy in the field. An early low-resolution cryo-EM structure of the yeast PIC in complex with eIF1 and eIF1A lacked a contact between 40S h18 and h34 at the mRNA entry channel, termed the 'latch', while 40S•eIF1A complexes and 40S alone displayed a closed latch, indicating a role for eIF1 in promoting latch opening (127). However, subsequent structures of 40S•eIF1•eIF1A complexes display a closed latch, calling this interpretation into question (138, 195). A recent cryo-EM structure of the 48S PIC by Llacer, et al. (included in Chapter 2 of this work) assembled on mRNA containing an AUC rather than an AUG start codon and including WT eIF2 and Met-tRNA$_i$ rather than variants intended to stabilize the closed complex resulted in a structure of the PIC in which the latch is open (py48S-open) (111). A second 48S PIC structure containing an mRNA with an AUG codon, presumably representing a state just after AUG recognition but before eIF1 release, demonstrates a closed latch (py48S-closed). Comparison of these two structures reveals an upward movement of the 40S head from the body in py48S-open compared to py48S-closed, resulting in a widened entry channel lacking fixed contacts to mRNA and an incompletely formed P site that are expected to render the open complex conducive to mRNA scanning (117). Thus, it is believed that these two structures provide us with "snapshots" of the PIC in its scanning conformation and when arrested at the start codon, and some of the structural rearrangements that accompany the transition between the two states (Fig. 5.1.1).

The py48S-open and py48S-closed complexes also resolved the structure and position of eIF2β within the PIC for the first time, revealing its contacts with eIF1 and eIF1A in the open PIC, as well as differential contacts with Met-tRNA$_i$ in the open and closed complexes. Through genetic and biochemical analyses described in Chapter 3, we have demonstrated that eIF2β contacts with eIF1, which are specific to the open complex, play a functional role in stabilizing eIF1 association with the open PIC, preventing its release during scanning before an appropriate start codon is recognized. Substitutions designed to disrupt the eIF2β:eIF1 interface resulted in hypoaccuracy (Sui$^-$) phenotypes and TC loading defects (Gcd$^-$) *in vivo*, and stabilized PIC complexes *in vitro*, indicating that they promote transition to the closed state. These data are consistent with destabilization of the open complex, as would be expected for abrogation of contacts that stabilize eIF1 binding. In addition, the role of eIF1 in accelerating TC binding to the open complex has not been well understood in molecular terms. While it is likely that eIF1 promotes the open state, at least in part, through clashes with Met-tRNA$_i$ in the closed PIC, contacts between eIF1 and eIF2β could play a role in facilitating TC loading onto the open PIC. The Gcd$^-$ phenotypes reported here for substitutions that impede interaction of eIF2β with eIF1 support the hypothesis that eIF1 is involved in recruitment of TC to the PIC through its contacts with eIF2β.

Additionally, we have elucidated two roles for the contacts of eIF2β with Met-tRNA$_i$. In the py48S-open structure, the P site is widened, and contacts between Met-tRNA$_i$ and the 40S body are not formed. Instead, eIF2β binds the D loop and anticodon stem loop, and

our data demonstrate that these contacts help to accelerate the recruitment of Met-tRNA$_i$ to the PIC and stabilize its binding in this position within the open complex. Substitutions designed to abrogate eIF2β contacts with the ASL in the open state resulted in increased initiation at UUG codons (Sui$^-$) and an accompanying defect in TC loading (Gcd$^-$) *in vivo,* indicative of destabilization of the open state. *In vitro,* one such substitution decreased the TC on-rate, consistent with the Gcd$^-$ phenotype, while impairing the TC off-rate. These apparently contradictory phenotypes are, in fact, consistent with the expected destabilization of the closed state, which slows initial TC loading while promoting transition to the closed state in which TC is bound more stably. In addition, comparison of the py48S-open and py48S-closed structures reveals clashes predicted between eIF2β in its open state position and Met-tRNA$_i$ in its closed state position. Substitutions designed to alleviate a clash of eIF2β with the D loop decreased the stringency of start codon selection *in vivo* and stabilized PICs *in vitro,* indicating that they facilitate transition to the closed state. Notably, none of these substitutions displayed a defect in TC loading, consistent with our interpretation that they remove an impediment rather than by destabilizing the open state. Our findings support the conclusion that eIF2β hinders the transition of Met-tRNA$_i$ to the closed state in the absence of an appropriate start codon through steric and electrostatic clashes. Clashes of eIF1 Loops 1 and 2 with Met-tRNA$_i$ in the closed PIC were reported to prevent premature transition to the closed state at non-AUG codons in a manner analogous to the clash between Met-tRNA$_i$ and eIF2β reported here. The contacts between eIF2β and eIF1 discussed above raise the possibility that these two factors work in concert to impede Met-tRNA$_i$ transition to its closed state position. That the rearrangement of eIF2β upon start codon recognition

alleviates its clash with Met-tRNA$_i$ and simultaneously abrogates its contacts with eIF1, allowing eIF1 displacement, supports this hypothesis. Thus, we find that eIF2β plays three distinct roles in start codon selection: 1) it stabilizes the binding of eIF1 on the open PIC, promoting the open state, 2) it stabilizes binding of Met-tRNA$_i$ in the open complex, and 3) it prevents transition to the closed PIC at non-AUG codons through clashes with the closed state position of Met-tRNA$_i$ (Fig. 5.1.2). Like eIF1 and the eIF1A CTT, we find that eIF2β has a dual function, promoting the open state of the PIC while preventing transition to the closed state prior to AUG recognition (Fig. 5.1.3) (302).

Although it has long been known that the sequence context around a start codon is critical to determining whether it will be utilized by the scanning ribosome, the molecular mechanism of context recognition has remained a mystery. Because of its position in the mRNA exit channel and its contacts with the important -3 context nucleotide in the mRNA, we examined the role of Rps26/eS26 in this process. Our results, described in Chapter 4, indicate that Rps26 does play a nuanced role in the recognition of start codon context, and demonstrate a decoupling of start codon selection from context recognition that may shed light on this mechanism. A purine:purine stacking interaction of the 18S rRNA with the -3 nucleotide, which could delay scanning and promote start codon recognition with a -3 purine base, has not been fully investigated. Nevertheless, it appears increasingly likely that the context nucleotides in the E-site and exit channel are recognized not by a single, decisive interaction, but by a network of contacts, influenced by the shape of the exit channel itself. Future work should examine this possibility, as well as the contributions of rRNA and other exit channel ribosomal proteins to context

recognition. In examining contacts between Rps26/eS26 and mRNA further upstream in the exit channel, we discovered that three well conserved basic Rps26 residues, previously implicated in crosslinking to mRNA, were important for stabilization of the closed PIC, presumably, by electrostatic attractions with the phosphodiester backbone of the mRNA (Fig. 5.1.2). Basic residues of Rps3 in the entry channel were implicated in a similar mechanism, suggesting a model in which ribosomal proteins at the entry and exit channels, possibly in concert with eIF3, clamp mRNA following start codon selection, stabilizing the closed PIC and aiding in the cessation of scanning (119). Further research is needed to examine this model, and the role eIF3 may play in this process. Because eIF3 is present at both the entry and exit channels, it would be interesting to determine whether it may act as a link between the stabilizing contacts of ribosomal proteins with mRNA upstream and downstream of the start codon.

Data from each of these studies inform our understanding of translation initiation. Thus, we can modify our model of start codon selection as follows (12, 112, 117, 194, 302). eIF1 and eIF1A, either alone or in a multi-factor complex with eIF3 and eIF5, bind to the 40S subunit to form a pre-initiation complex, promoting an open conformation in which the mRNA entry channel and P site are widened (Fig. 5.1.4 (i)). Interactions between members of the MFC and Met-tRNA$_i$ facilitate recruitment of TC to the PIC, where it is stabilized in the open conformation in part by contacts between eIF2β and Met-tRNA$_i$ and eIF2β with eIF1, promoting continued scanning and increasing the stringency of AUG selection. Interactions of eIF4G with 40S-bound eIF3 or eIF5, along with the helicase activity of eIF4A, in complex with cap-bound eIF4F and stimulated by eIF4B,

serve to recruit the open PIC to the 5' end of mRNA. The open conformation of the PIC facilitates scanning in the 5'-to-3' direction assisted by the helicases Ded1 and eIF4A, and enhanced by eIF4B, and the contacts between eIF2β and eIF1 help to anchor eIF1 onto the complex and prevent its premature release at non-AUG codons (Fig. 5.1.4 (ii)). When an appropriate start codon is encountered, base pairing with the anticodon of Met-$tRNA_i$ and auxiliary interactions of the context nucleotides with Rps26/eS26, as well as eIF2α, Rps5/uS7, and other exit channel components, ensure its recognition. Start codon recognition triggers a conformational change in the 40S ribosome that results in head closure, including closure of the latch, a constricted entry channel, and a compact P site that fully encases Met-$tRNA_i$ and strongly stabilizes its association with the PIC. These conformational changes are attended by clashes between eIF2β and Met-$tRNA_i$ (Fig. 5.1.4 (iii)). Movement of eIF2β away from the Met-$tRNA_i$ ASL and eIF1 facilitates this deeper binding of Met-$tRNA_i$ in the P site, and the resulting clashes between Met-$tRNA_i$ and eIF1 Loops 1 and 2, as well as the loss of eIF1:eIF2β contacts, allow displacement of eIF1 on the 40S platform (Fig. 5.1.4 (iv)), followed by its dissociation from the 40S subunit. Met-$tRNA_i$ in the closed complex also clashes with the eIF1A-CTT, which dissociates from the eIF5-CTD and moves from the P site to interact with the eIF5-NTD, which is now bound to the platform in place of eIF1. The eIF5-NTD directly stabilizes Met-$tRNA_i$ binding in the P site in a position and orientation closer to that achieved following 60S subunit joining; and the eIF5-CTD is now free to interact with eIF2β, further stabilizing the closed state. Interactions of the mRNA with Rps3/uS5 in the entry channel and Rps26/eS26 in the exit channel also help to stabilize the closed complex and prevent further ribosomal scanning. The newly formed interaction between the eIF1A-

CTT and the eIF5-NTD gates $P_i$ release. The resultant completion of GTP hydrolysis on eIF2 and movement of the eIF5-NTD to the eIF1 binding site, which prevents reassociation of eIF1, making start codon selection irreversible (Fig. 5.1.4 (v)). eIF2•GDP then dissociates from the PIC in complex with eIF5 and the eIF1A-CTT recruits eIF5B, which facilitates joining of the large ribosomal subunit. The GTP on eIF5B is hydrolyzed, triggering release of eIF5B and eIF1A. Ultimately, translation initiation ends with Met-$tRNA_i$ base paired to a start codon in the P site of an 80S ribosome bound by eIF3 but lacking other initiation factors, prepared to bind an elongator tRNA in the A site and proceed into elongation. Future research into this complex molecular mechanism will continue to illuminate the order of events after the scanning PIC encounters an appropriate start codon, and how these interactions come together to determine the frequency with which a near cognate codon will be utilized or an AUG skipped over.

It is increasingly clear that the ribosome functions through a complex network of molecular interactions between rRNA, ribosomal proteins, and initiation factors. These interactions do not only ensure accurate recognition of AUG codons, but they optimize the fidelity of start codon selection. Contacts that preferentially stabilize the open conformation or impede transition to the closed state restrict initiation at non-AUG codons or AUGs in poor sequence context by promoting continued scanning past these suboptimal initiation sites. By contrast, interactions that specifically stabilize the closed state are crucial for infrequent but necessary recognition of these suboptimal start codons. Presumably, evolution has tuned the relative strengths of these opposing interactions to

set the initiation probabilities at poor versus strong initiation sites in order to ensure optimum expression of the translatome. The data described here add new interactions to this network and improve our understanding of the precise molecular mechanism and accuracy of start codon selection. No doubt, future research will uncover yet more interactions that play a role in stipulating the carefully tuned fidelity of start codon selection. An understanding of this mechanism will further efforts to decode the expression of proteins in healthy cells, diseased cells, and developing cells.

**5.1 Figures and Tables**

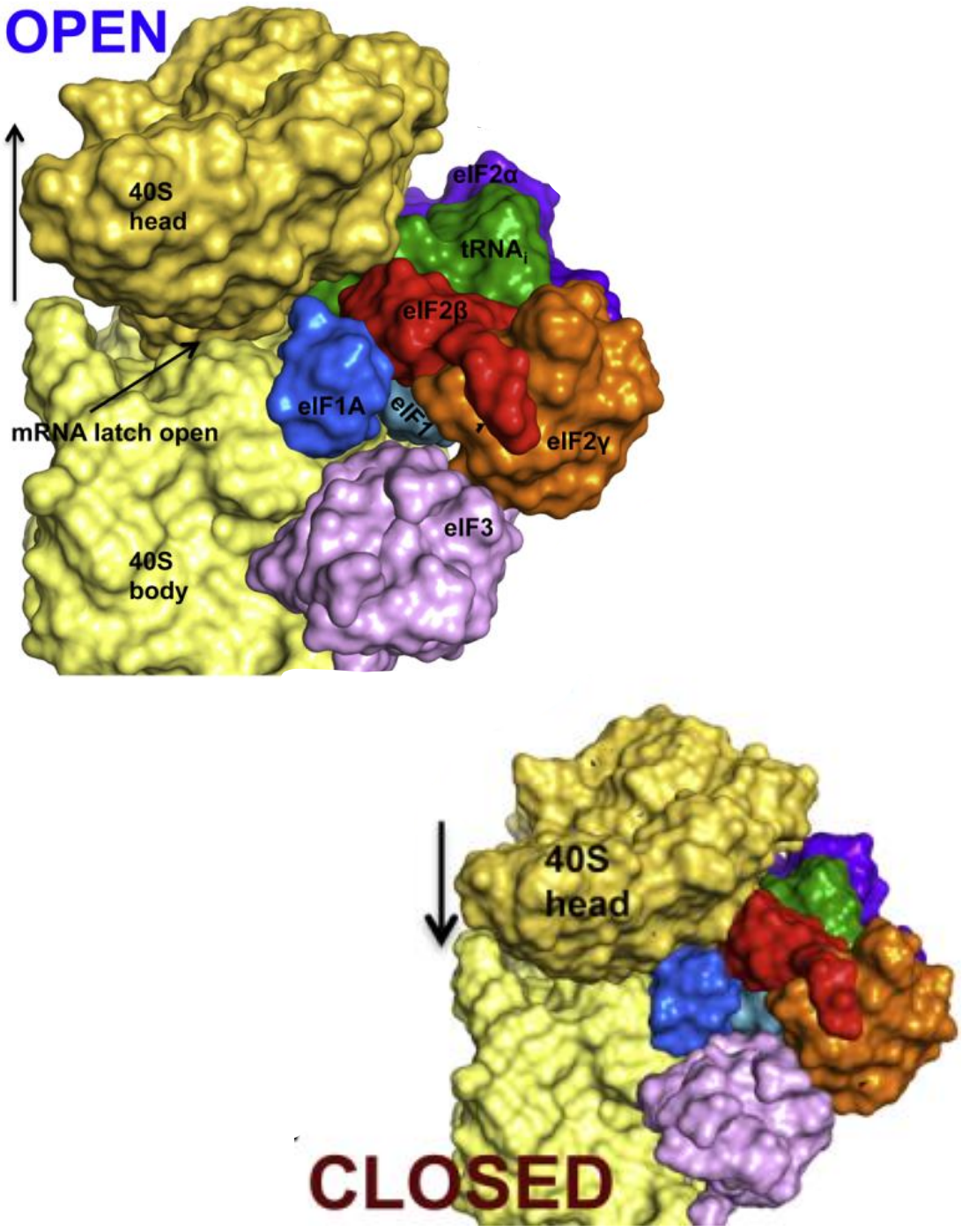
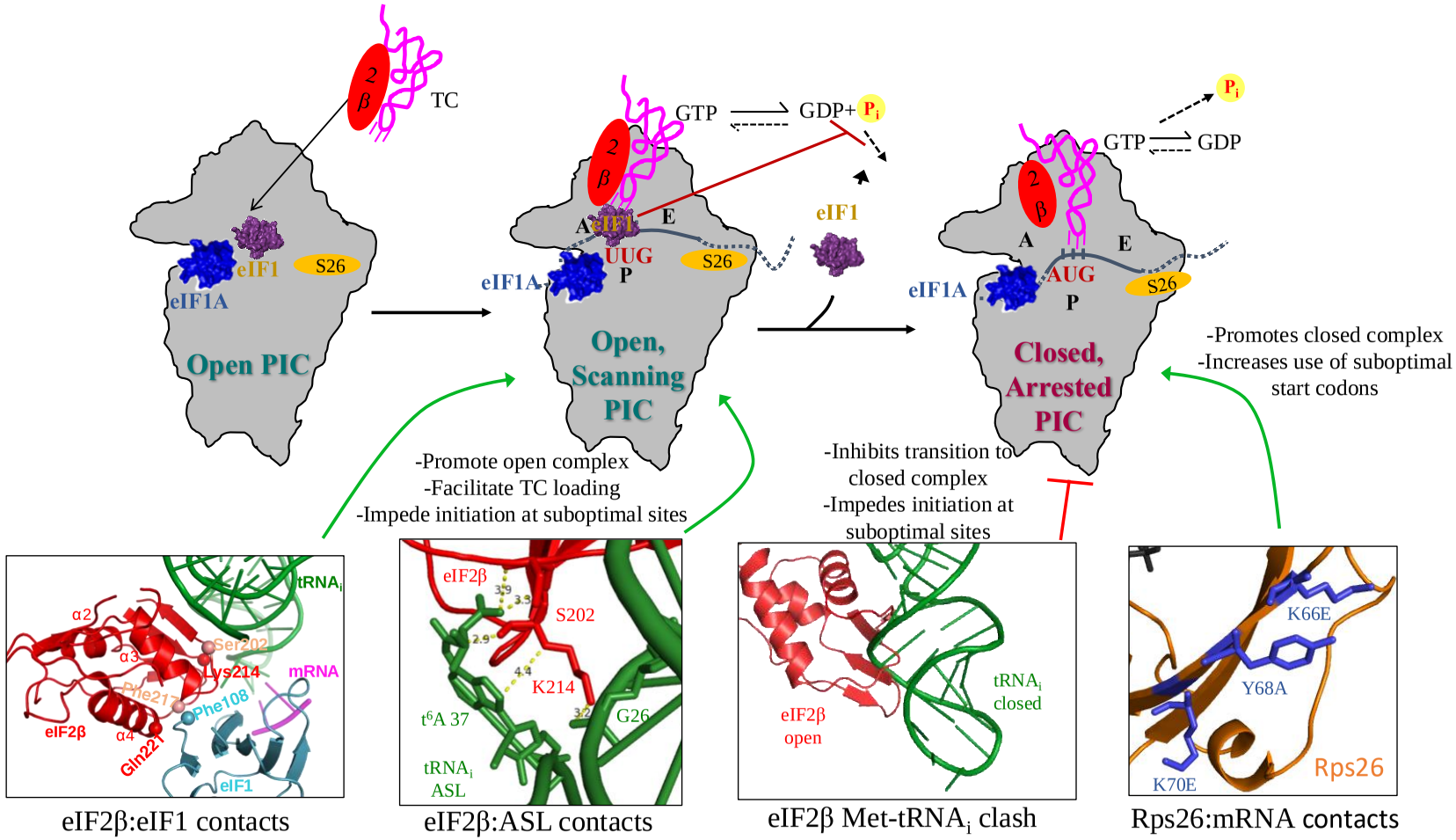


**Figure 5.1.1. Cryo-EM structures of yeast partial 48S initiation complexes in open and closed states reveal insights into ribosomal scanning and start codon recognition.**

The py48S-open structure displays a widened mRNA entry channel, open latch, and partially formed P site in which many contacts between Met-tRNA$_i$ and the 40S body are absent. This structure is presumably conducive to scanning, allowing movement of the mRNA and mRNA sampling by Met-tRNA$_i$. In the py48S-closed structure, the 40S head is repositioned downward relative to the body, constricting the mRNA entry channel, and the latch is closed. These changes likely restrict mRNA movement and assist with the cessation of scanning upon AUG recognition. In addition, Met-tRNA$_i$ is bound more deeply in the P site ($P_{IN}$), facilitating its base pairing with the AUG start codon, and the P site itself is constricted, stabilizing Met-tRNA$_i$ in the $P_{IN}$ state. (PDB 3JAQ, 3JAP)




eIF2β:eIF1 contacts | eIF2β:ASL contacts | eIF2β Met-tRNA$_i$ clash | Rps26:mRNA contacts

**Figure 5.1.2. A network of interactions promotes either the open or the closed PIC.**

The interactions investigated in this work join a growing network of molecular interactions that tune the accuracy of start codon selection. Contacts of eIF2β with eIF1 and the ASL of Met-tRNA$_i$ stabilize the open, scanning conformation of the PIC. As a result, these interactions facilitate TC loading and impede initiation at suboptimal start codons, such as near cognate codons or AUG codons in poor sequence context, increasing initiation fidelity. Clashes between eIF2β and the D loop of Met-tRNA$_i$ block transition to the closed state but do not contribute to stabilization of the open state, hindering initiation at suboptimal start codons without impacting TC loading. Finally, contacts between Rps26/eS26 and mRNA context nucleotides in the exit channel promote the closed state. These interactions encourage initiation at suboptimal start codons, decreasing initiation accuracy. The balance of interactions that bolster or restrict initiation accuracy, along with the placement, identity, and context of start codons, has likely been tuned through evolution to hone the expression levels of cellular proteins.

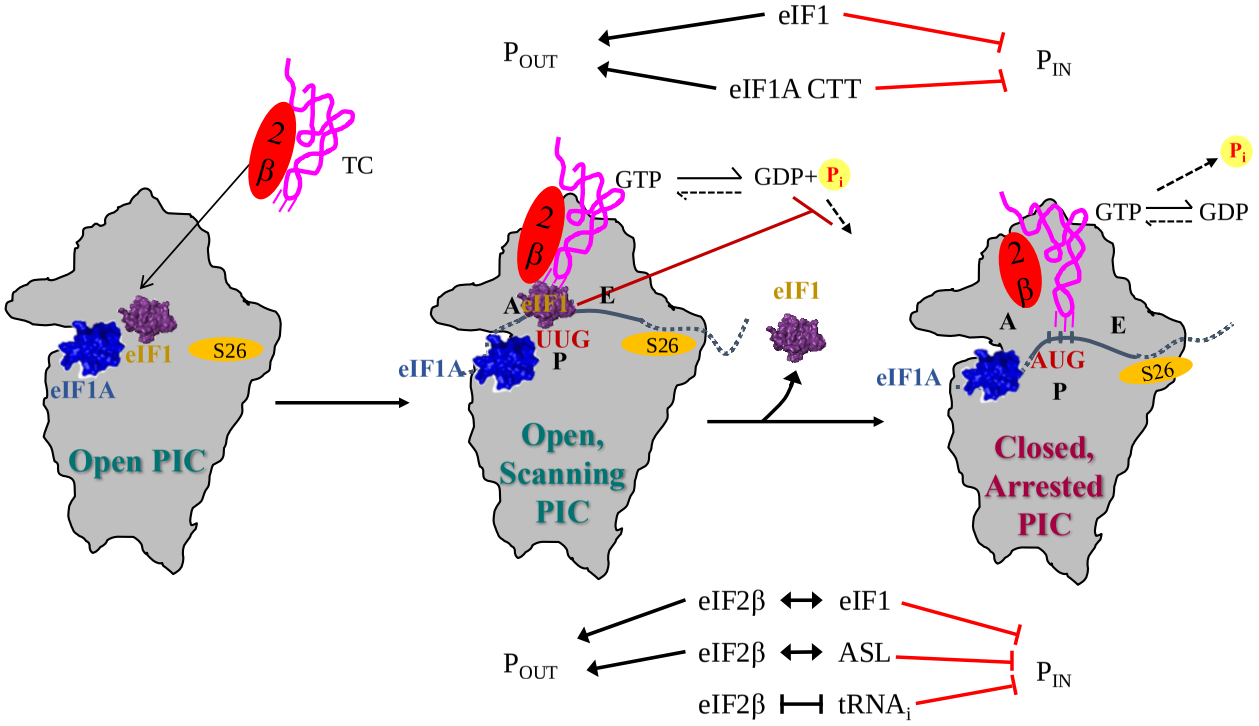

eIF1
P_OUT
P_IN
eIF1A CTT
2
β
TC
GTP
GDP+ P_i
P_i
GTP
GDP
eIF1
S26
eIF1A
Open PIC
A
E
P
UUG
AUG
Open, Scanning PIC
Closed, Arrested PIC
eIF2β
ASL
tRNA_i

**Figure 5.1.3. Multiple initiation factors play a dual role in start codon selection.**

It has been previously reported that both eIF1 and the CTT of eIF1A play a dual role in translation initiation, promoting the open state of the PIC, while simultaneously inhibiting transition to the closed state – in both cases by occupying the P site and sterically obstructing movement of Met-tRNA$_i$ into its closed state position. In this work, we find that eIF2β similarly fulfills a dual role. Interactions of eIF2β with eIF1 and the ASL of Met-tRNA$_i$ stabilize the open, scanning conformation of the PIC. As a result, like eIF1 and the eIF1A CTT, these contacts of eIF2β stimulate TC loading and increase the stringency of start codon fidelity by encouraging continued scanning past suboptimal start codons. Mutations that abrogate these contacts result in Sui$^-$ and Gcd$^-$ phenotypes similar to those seen in mutations of eIF1 that impair binding of eIF1 to the 40S subunit. At the same time, steric and electrostatic clashes between eIF2β in its open state position and the closed state position of Met-tRNA$_i$ impede rearrangement of the PIC to its closed conformation, hampering transition to the closed state. These interactions also work to restrict initiation to AUG codons in good context, although they do not affect binding of TC to the PIC because they do not stabilize the open state onto which TC loads. Like mutations in eIF1 Loop 2, substitutions in eIF2β that alleviate a clash produce Sui$^-$ phenotypes without an accompanying Gcd$^-$ phenotype. Thus, we see a trend in the function of initiation factors involved in start codon selection to simultaneously form contacts that stabilize the open PIC while sterically or electrostatically occluding transition to the closed state. This may represent an evolutionary strategy to manage the complexity and importance of correctly tuning the accuracy of start codon recognition given the limited space available for regulating factors in the vicinity of the P site.

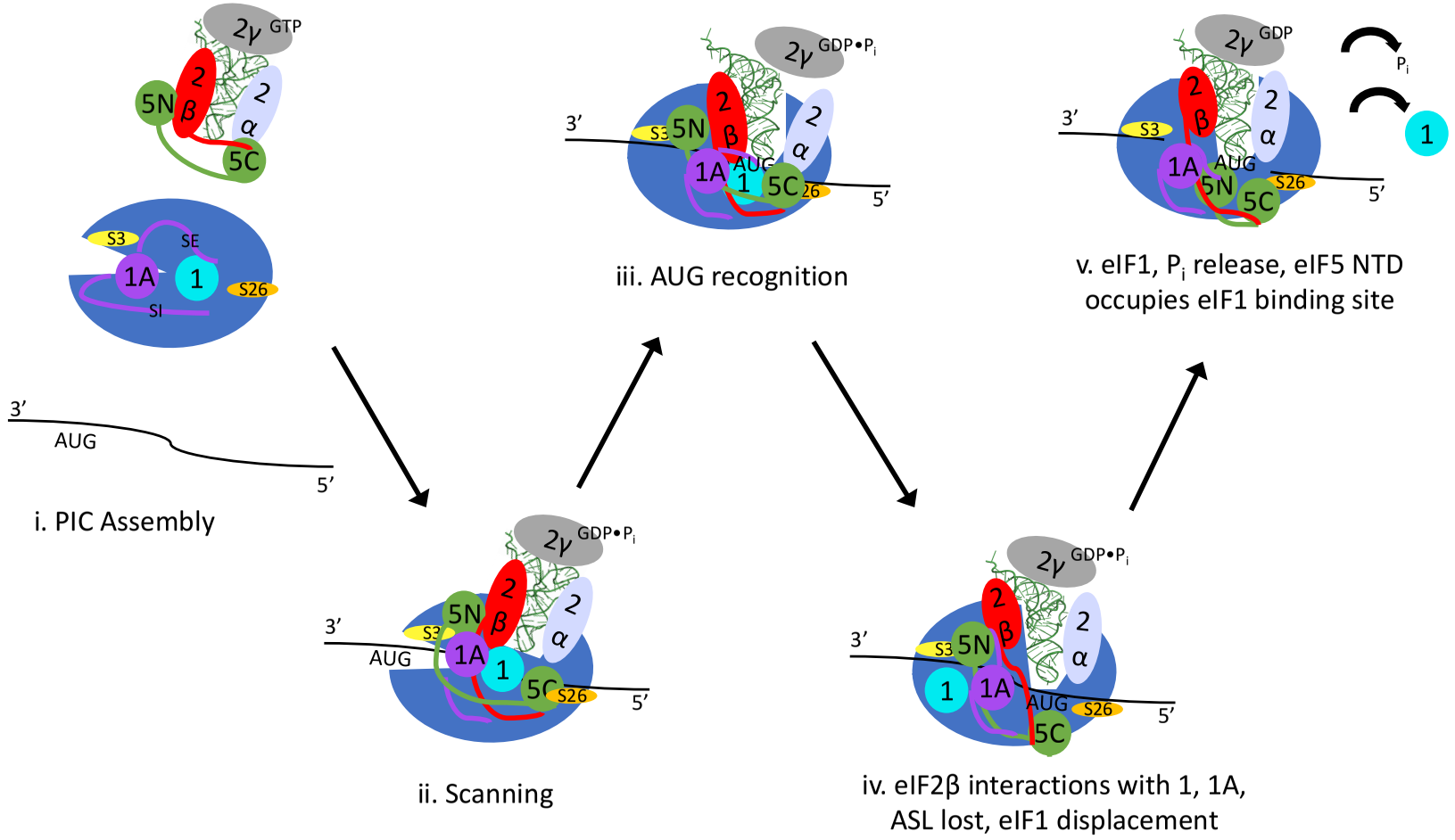

2γ GTP
2
β
2
α
5N
5C
S3
SE
1A
1
S26
SI
3′
AUG
5′
i. PIC Assembly
2γ GDP•Pi
3′
AUG
5′
ii. Scanning
2γ GDP•Pi
3′
AUG
5′
iii. AUG recognition
2γ GDP•Pi
3′
AUG
5′
iv. eIF2β interactions with 1, 1A, ASL lost, eIF1 displacement
2γ GDP
Pi
1
3′
AUG
5′
v. eIF1, Pi release, eIF5 NTD occupies eIF1 binding site

**Figure 5.1.4. New model of start codon recognition**

Based on the data presented in this work, we can modify the model of start codon selection presented in Fig. 1.7.7 as follows. (i) Binding of eIF1 and the scanning enhancer elements of the eIF1A CTT to the 40S subunit promote the open conformation, facilitating binding of TC and mRNA loading. (ii) The resulting open PIC is conducive to scanning, during which GTP is hydrolyzed to GDP•$P_i$, but release of inorganic phosphate is blocked by eIF1. eIF2β promotes the open state and continued scanning through contacts with eIF1 that stabilize eIF1 binding on the ribosome and interactions with the Met-$tRNA_i$ ASL that stabilize it in the open PIC. (iii) When the scanning PIC encounters an appropriate start codon, it is recognized through base pairing to the anticodon of Met-$tRNA_i$, as well as interactions of Rps26/eS26, Rps5/uS7, eIF2α, and other E site and exit channel components with the upstream sequence context. These interactions promote transition to the closed state in response to an AUG codon in good context and aid in the cessation of scanning. Clashes with several factors in the P site, including eIF1, eIF2β, and the CTT of eIF1A, impose an energetic barrier to the movement of Met-$tRNA_i$ deeper into the P site necessitated by base pairing with the start codon, requiring correct base pairing to overcome this impediment. (iv) During transition to the closed state, eIF2β moves toward eIF2γ, becoming more compact and losing its contacts with eIF1, eIF1A, and the ASL. In the process, clashes of eIF2β with the Met-$tRNA_i$ ASL and D loop are alleviated. The resultant deeper binding of Met-$tRNA_i$ in the P site results in clashes with Loops 1 and 2 of eIF1. In response to these clashes and the loss of contacts with eIF2β, eIF1 is displaced from the P site to a secondary binding site on the 40S platform, while the eIF1A CTT moves out of the P site and interacts with the eIF5 NTD.

In the absence of the eIF1A CTT, the CTD of eIF5 forms a strong interaction with the eIF2β NTT. (v) eIF1 dissociates from the PIC, and its release, along with the interaction of the eIF1A CTT with the eIF5 NTD, gates $P_i$ release from eIF2. The eIF5 NTD moves to occupy the eIF1 binding site, stabilizing Met-$tRNA_i$ in the closed state and preventing rebinding of eIF1. Interactions between the mRNA and Rps3/us5 in the entry channel and Rps26/eS26 in the exit channel also help to stabilize the closed PIC and prevent further ribosomal scanning.

# 6. Appendix A: Supplemental Material for "Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex"

Supplemental Material for Chapter 2:

Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex

Jose L. Llácer,[1,4] Tanweer Hussain,[1,4] Laura Marler,[2] Colin Echeverría Aitken,[3] Anil Thakur,[2] Jon R. Lorsch,[3,5] Alan G. Hinnebusch,[2,5] and V. Ramakrishnan[1,*]

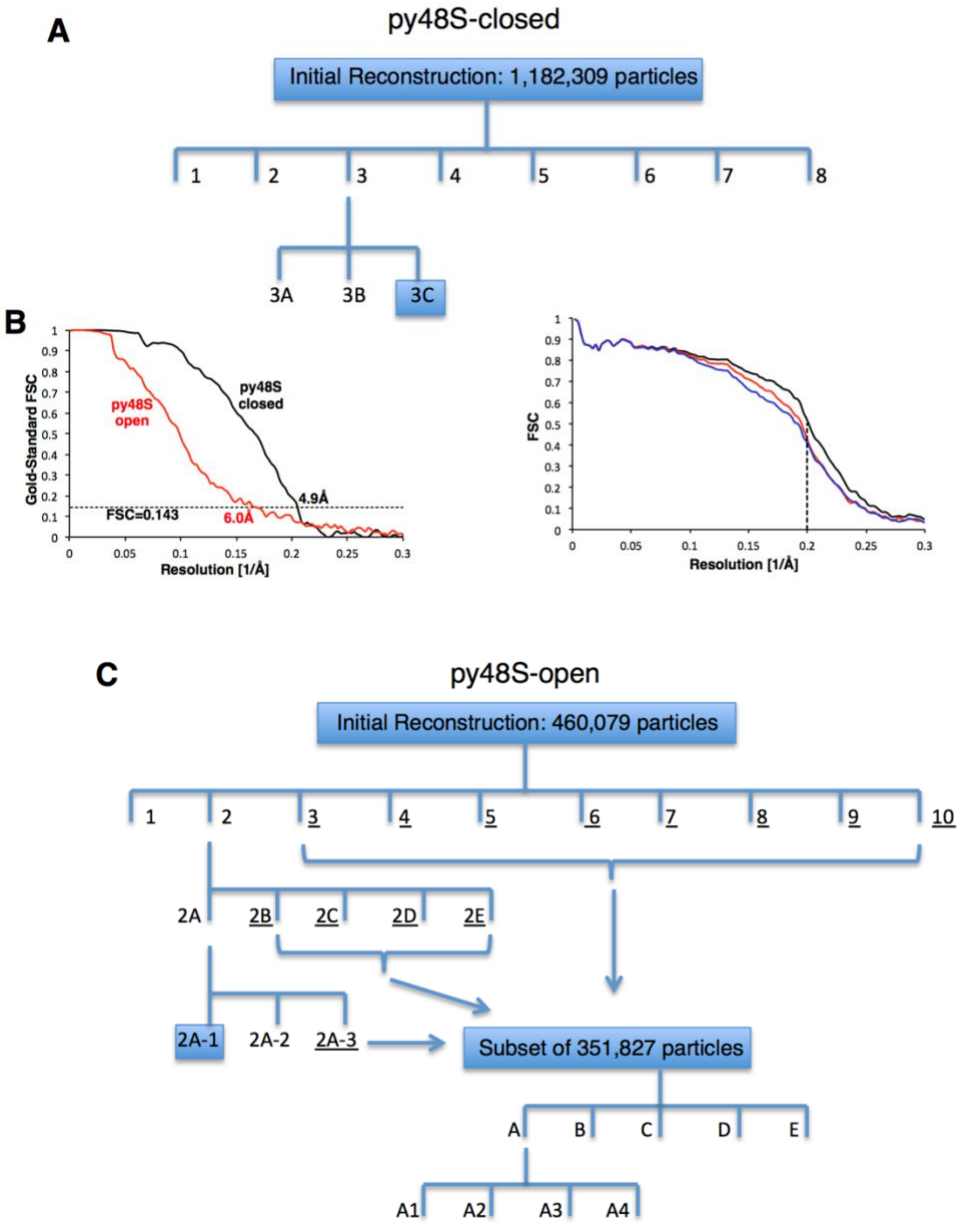
Mol Cell. 2015 Aug 6; 59(3): 399–412.

**Molecular Cell, Volume *59* Supplemental Information**

# Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex

**Jose L. Llácer, Tanweer Hussain, Laura Marler, Colin Echeverría Aitken, Anil Thakur, Jon R. Lorsch, Alan G. Hinnebusch, and V. Ramakrishnan**

## 6.1 SUPPLEMENTARY FIGURES AND TABLES



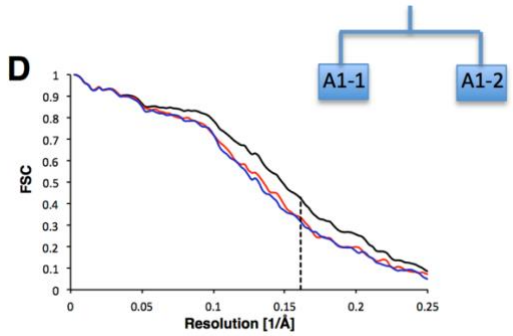



**Figure 6.1.1. Maximum-likelihood 3D classification scheme, related to Figure 2.10.1**

Maximum-likelihood 3D classification schemes (See Supplemental Experimental Procedures): (A) py48S-closed complex: Class 3C (21,401 particles; 4.9 Å) corresponding to py48S-closed complex is highlighted in a box.

(B) At the left, Gold-standard Fourier Shell Correlation (FSC) curves for the py48S-closed (black) and py48S-open (red) complexes. At the right, analysis of overfitting by cross-validation of the py48S- closed model. $FSC_{work}$ curves (red) corresponding to the refined model versus the half-map it was refined against, and $FSC_{test}$ curves (blue), i.e. those calculated between the refined atomic model and the other half-map. The black curve shows the FSC curve between a reconstruction from all particles and the model refined against the map. The dashed line represents the highest resolution (5.0Å) used in these refinements.

(C) py48S-open complex: Class 2A-1 (86,055 particles; 3.5 Å) corresponding to 40S eIF1 eIF1A complex, Class A1-1 (4,547 particles; 6.0 Å) corresponding to py48S-open complex and Class A1-2 (1,580 particles; 14.9 Å) corresponding to py43S are highlighted

in a box. (D) Analysis of overfitting by cross-validation of the py48S-open model. $FSC_{work}$ curves (red) corresponding to the refined model versus the half-map it was refined against, and $FSC_{test}$ curves (blue), i.e. those calculated between the refined atomic model and the other half-map. The black curve shows the FSC curve between a reconstruction from all particles and the model refined against the map. The dashed line represents the highest resolution (6.2Å) used in these refinements.

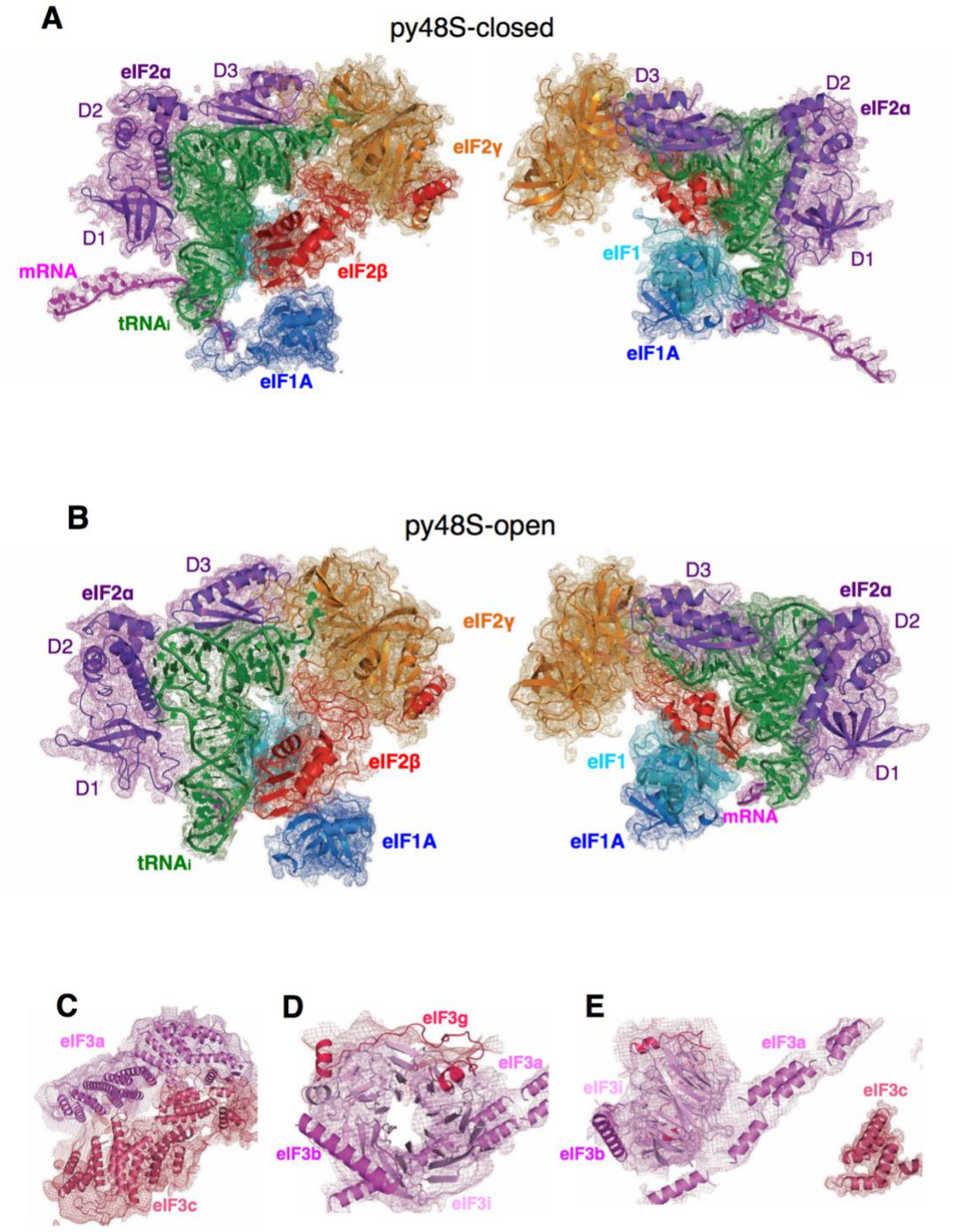
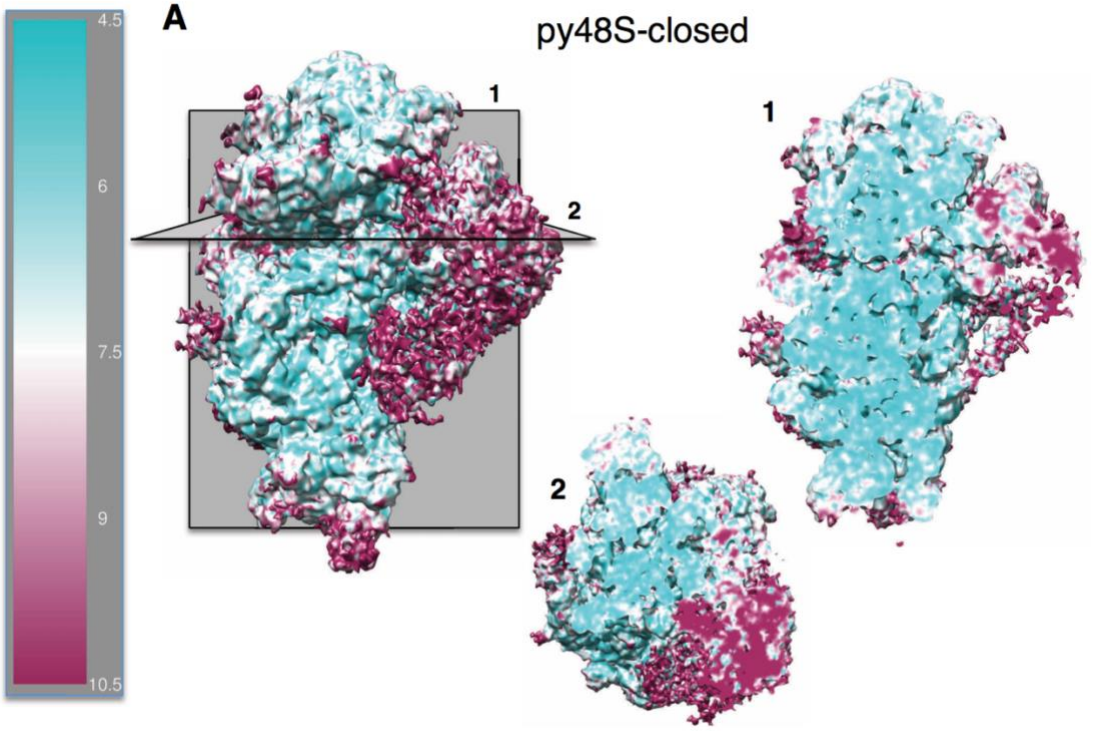

A
py48S-closed
eIF2α
D3
D2
eIF2γ
D1
mRNA
eIF2β
tRNAi
eIF1A
D3
D2
eIF2α
eIF1
D1
eIF1A
B
py48S-open
D3
eIF2α
eIF2γ
D2
eIF2β
D1
eIF1A
tRNAi
D3
eIF2α
D2
eIF1
D1
mRNA
eIF1A
C
eIF3a
eIF3c
D
eIF3g
eIF3a
eIF3b
eIF3i
E
eIF3a
eIF3i
eIF3c
eIF3b

**Figure 6.1.2. Fitting of ligands in density maps, related to Figure 2.10.1**

(A) Fitting of eIF1, eIF1A, eIF2, $tRNA_i$ and mRNA in py48S-closed map at 4.9 Å shown in two orientations. (B) Fitting of eIF1, eIF1A, eIF2, $tRNA_i$ and mRNA in py48S-open map at 6.0 Å shown in two orientations. (C) eIF3a/eIF3c PCI heterodimer in py48S-closed map. (D) eIF3b-CTD/eIF3i/eIF3g-NTD in py48S-closed map. (E) eIF3b-CTD/eIF3i/eIF3g-NTD trimer, cluster of eIF3c helices and bundle of eIF3a long helices in py48S-closed map.


A
py48S-closed
4.5
6
7.5
9
10.5
1
2
1
2

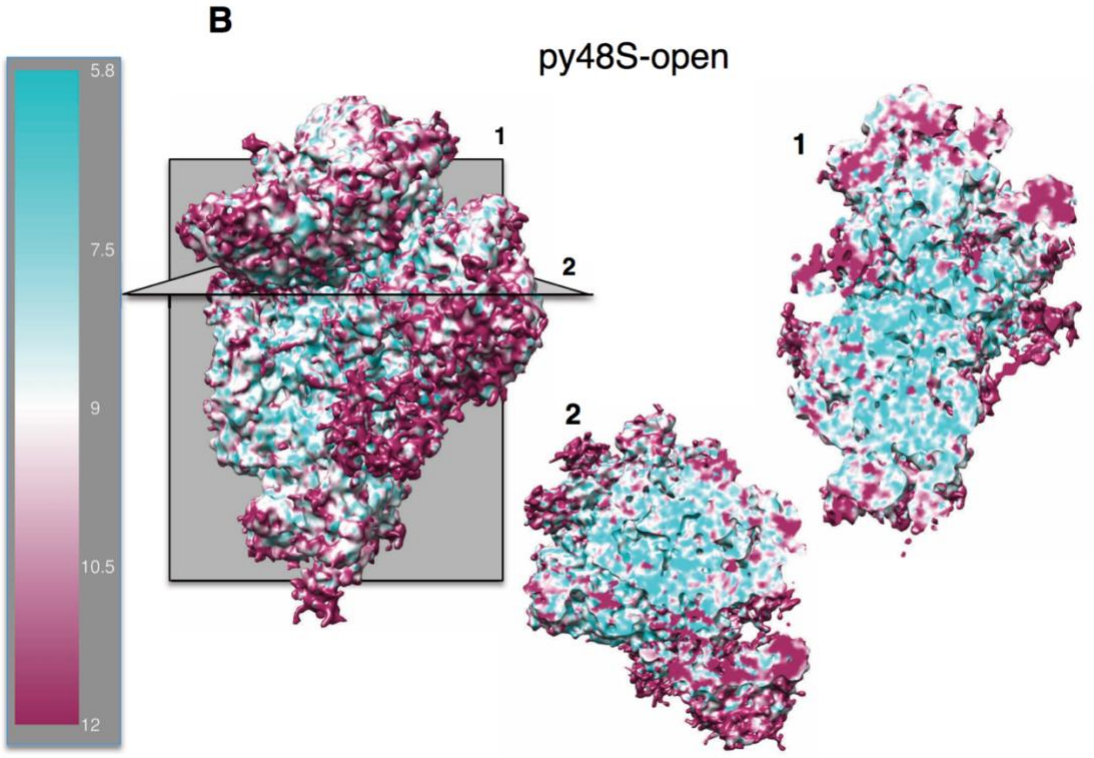

B
py48S-open
5.8
7.5
9
10.5
12
1
2
1
2

**Figure 6.1.3. Local resolution features, related to Figure 2.10.1**

(A) Surface (left) and cross-sections along the mRNA channel in two different planes of a 4.9 Å map, colored according to local resolution (See Experimental Procedures) of py48S-closed complex.

(B) Surface (left) and cross-sections along the mRNA channel in two different planes of a 6.0 Å map, colored according to local resolution (See Experimental Procedures) of py48S-open complex.

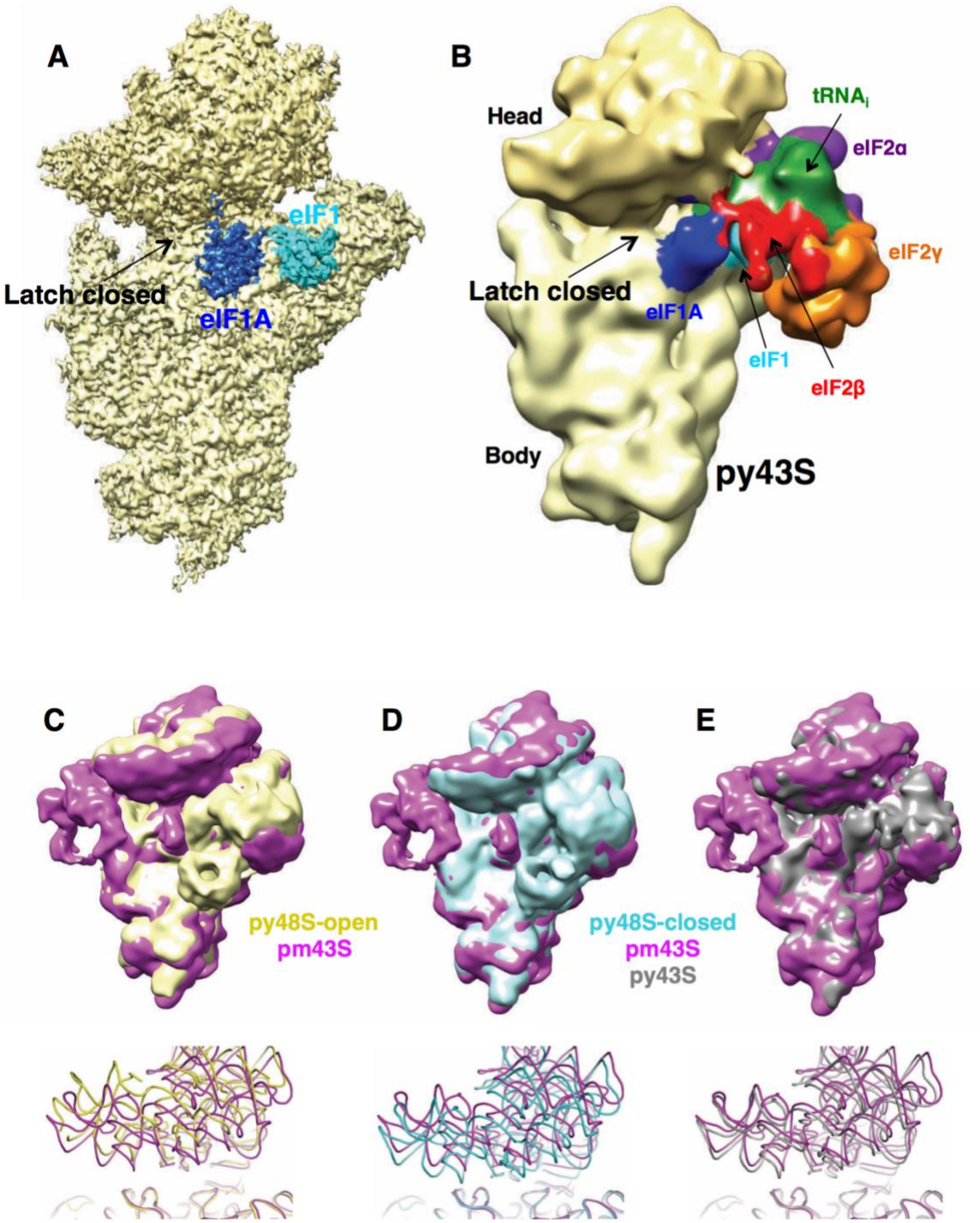
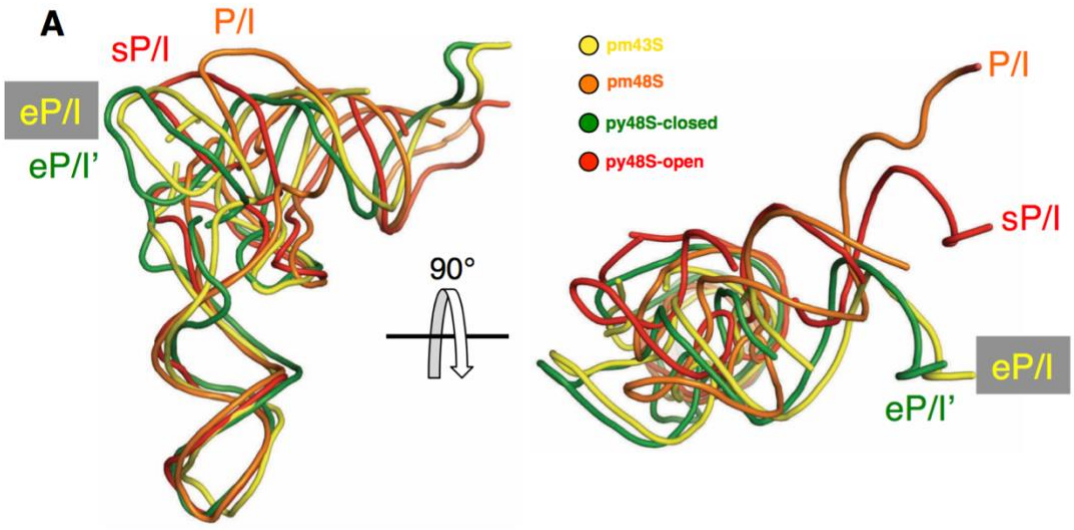

A
eIF1
Latch closed
eIF1A
B
Head
tRNAi
eIF2α
eIF2γ
Latch closed
eIF1A
eIF1
eIF2β
Body
py43S
C
py48S-open
pm43S
D
py48S-closed
pm43S
py43S
E

**Figure 6.1.4. Cryo-EM maps of py43S and 40S•eIF1•eIF1A complexes, related to Figures 2.10.1 and 2.10.2**

(A) Cryo-EM map of 40S□eIF1□eIF1A PIC at 3.5 Å. Density for eIF1 (cyan) and eIF1A (blue) can be clearly seen. The structure of the 40S□eIF1□eIF1A complex is similar to the PIC-2 complex reported earlier (Hussain et al., 2014) with an r.m.s.d. of 0.99 Å for 35,235 atoms of 18S rRNA.

(B) Cryo-EM map of py43S PIC at 15.0 Å. Density for eIF1, eIF1A and TC is observed. (C) Superimposition of py48S-open (yellow) and pm43S (magenta) (Hashem et al., 2013) maps. *Below:* Cartoon representation of rRNAs of the two structures. The head is clearly moved up in py48S-open. The py48S-open map is low-pass filtered to 12 Å. (D) Superimposition of py48S-closed (cyan) and pm43S (magenta) (Hashem et al., 2013) maps. *Below:* Cartoon representation of rRNAs of the two structures. The head is in a similar but not identical position in the two complexes. The py48S-closed map is low-pass filtered to 12 Å. (E) Superimposition of py43S (grey) and pm43S (magenta) (Hashem et al., 2013) maps. *Below:* Cartoon representation of rRNAs of the two structures. The head is almost identical in both complexes.


A
P/I
sP/I
eP/I
eP/I'
90°
pm43S
pm48S
py48S-closed
py48S-open
P/I
sP/I
eP/I
eP/I'

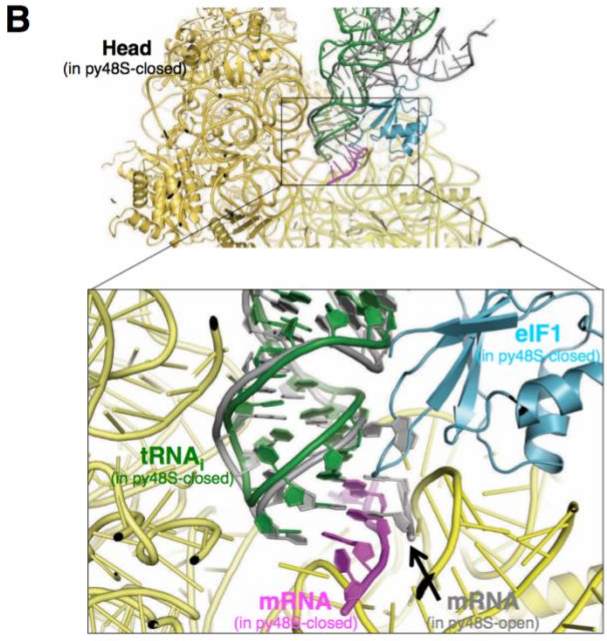

B
Head
(in py48S-closed)
eIF1
(in py48S-closed)
tRNAi
(in py48S-closed)
mRNA
(in py48S-closed)
mRNA
(in py48S-open)

**Figure 6.1.5. Distinct $tRNA_i$ conformations and the mRNA path, related to Figure 2.10.3**

(A) Two different views of the superimposition of $tRNA_i$ from different complexes aligned to the head of 40S. The conformation of $tRNA_i$ from py48S-closed (eP/I', green) is similar to that described for pm43S complex [eP/I, yellow; (Hashem et al., 2013)], but different from the P/I conformation (from pm48S PIC, in orange, from 4KZZ). The $tRNA_i$ from py48S-open (red) complex is in an orientation that appears closer to P/I than the eP/I. We have termed this orientation sP/I (scanning P/I). (B) Superimposition of py48S- open and closed complexes aligned to the 40S head shows the relative position of $tRNA_i$ in the P site. The body and head of py48S-closed complex is shown in yellow with its $tRNA_i$ in green. The mRNA (magenta) and β-hairpin 1 of eIF1 (cyan) of py48S-closed complex interact at the P site. The mRNA and $tRNA_i$ of py48S-open complex is shown in grey. The mRNA in the py48S- open complex would clash with the body of 40S in the py48S-closed complex.

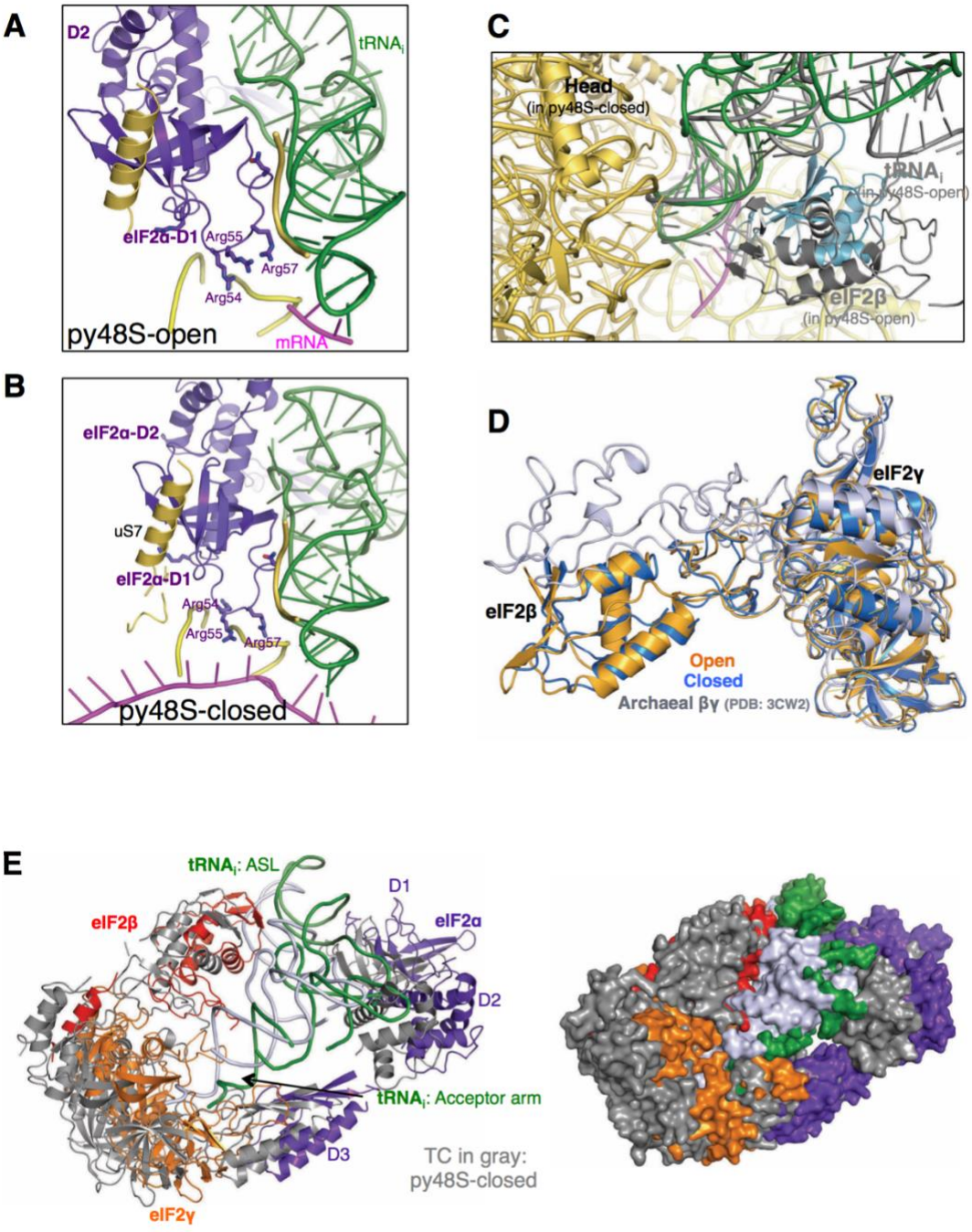

A
D2
tRNAi
eIF2α-D1
Arg55
Arg57
Arg54
py48S-open
mRNA
B
eIF2α-D2
uS7
eIF2α-D1
Arg54
Arg55
Arg57
py48S-closed
C
Head
(in py48S-closed)
tRNAi
(in py48S-open)
eIF2β
(in py48S-open)
D
eIF2γ
eIF2β
Open
Closed
Archaeal βγ (PDB: 3CW2)
E
tRNAi: ASL
D1
eIF2β
eIF2α
D2
tRNAi: Acceptor arm
D3
TC in gray:
py48S-closed
eIF2γ

**Figure 6.1.6. Relative orientations of eIF2 in py48S- open and closed PICs, related to Figure 2.10.5**

(A) eIF2α-D1 and ASL of $tRNA_i$ in py48S-open complex. The neighboring rRNA residues are shown in yellow lines and a helix of uS7 is also shown in yellow. Conserved arginines are shown. Arg55 and Arg57 project away from the modeled mRNA (grey; from py48S-closed complex), and Arg54, which in the closed complex interacts with the body of the 40S, comes closer to the mRNA. (B) eIF2α-D1 and the ASL of $tRNA_i$ in the py48S-closed complex. The neighboring rRNA residues are shown in yellow lines and a helix of uS7 is also shown in yellow. Conserved arginines are shown. Arg55 and Arg57 interact with the mRNA (magenta).

(C) Superimposition of py48S- closed and open complexes shows that eIF2β of py48S-open complex (grey) would clash with eIF1 (cyan) of py48S-closed complex, highlighting the need for the conformational change within the TC during the open to closed PIC transition. The 40S, $tRNA_i$, mRNA and eIF1 of the py48S-closed complex are shown in color while only $tRNA_i$ and eIF2β of py48S-open complex are shown in grey.

(D) Superimposition of eIF2-βγ dimer of py48S- open (yellow) and closed (blue) complexes with the most similar archaeal βγ dimer (grey, from 3CW2) using γ as a reference shows different position of the β subunit with respect to the γ subunit. (E) Relative position of TC in py48S-open (color) and py48S-closed (grey) based on superposition of the two complexes. The cartoon and surface representation show the relative position of each component of the TC.

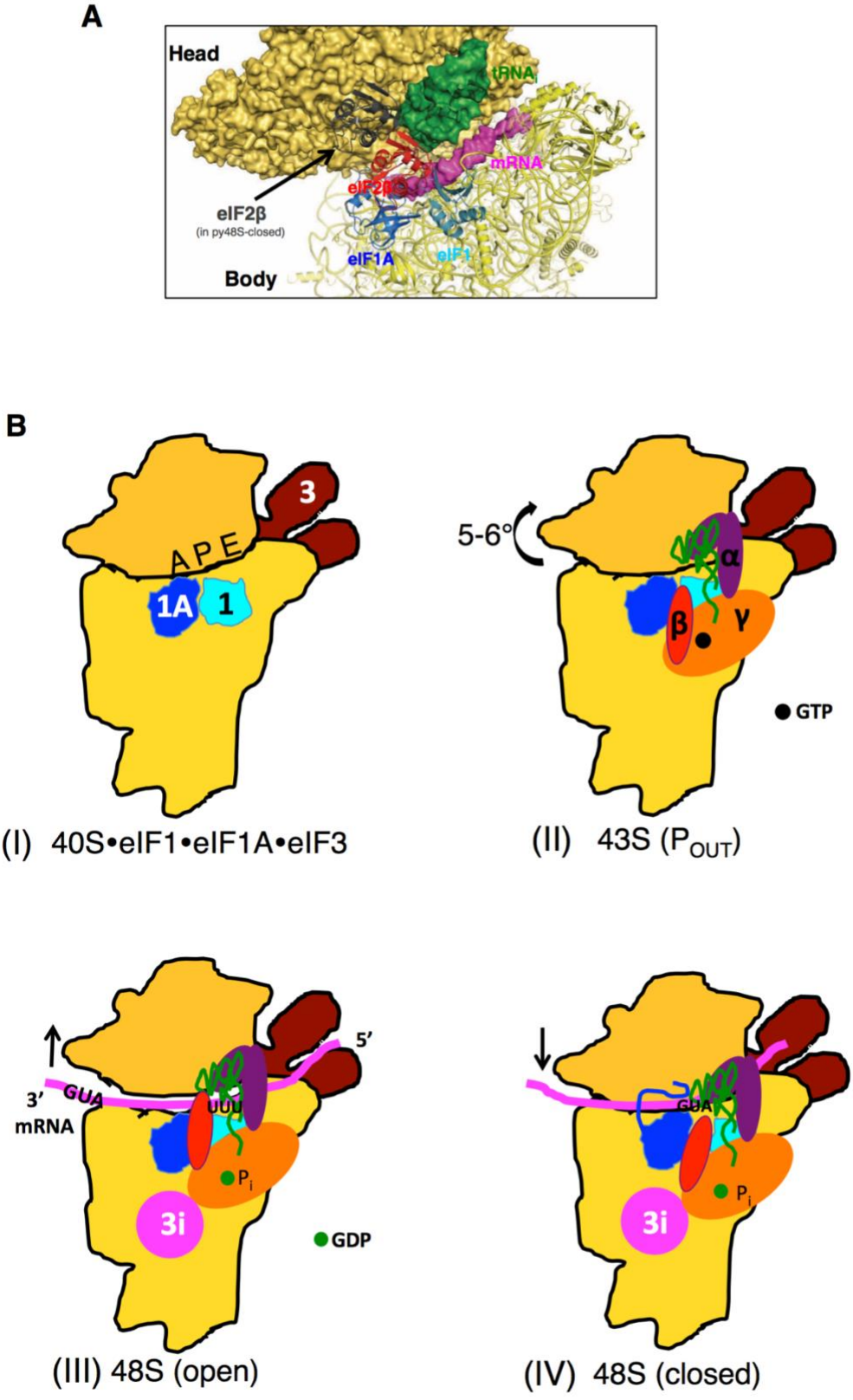

A
Head
tRNA
mRNA
eIF2β
eIF2β
(in py48S-closed)
eIF1A
eIF1
Body
B
3
A P E
1A
1
(I) 40S•eIF1•eIF1A•eIF3
5-6°
α
β
γ
GTP
(II) 43S (P$_{OUT}$)
5′
3′
mRNA
GUA
UUU
P$_i$
3i
GDP
(III) 48S (open)
GUA
P$_i$
3i
(IV) 48S (closed)

**Figure 6.1.7. Conformational changes from open to closed state, related to Figure 2.10.5**

(A) Opening of the mRNA channel. Modeling of the position of eIF2β (grey) (with respect to 40S head and $tRNA_i$) observed in the py48S-closed complex into the py48S-open complex shows how in this conformation the mRNA channel would be opened up. In the py48S-open complex, eIF2β (red) blocks mRNA access by forming interactions with $tRNA_i$ (green) attached to head and eIF1 (cyan) and eIF1A (blue) attached to the body. The acceptor arm of $tRNA_i$ is not shown for clarity.

(B) Major structural changes during eukaryotic translation initiation. Binding of eIF1, eIF1A and eIF3 to the 40S subunit (I) facilitates TC binding in the $P_{OUT}$ conformation to form the 43S PIC (II). Upward movement of the head expands the mRNA entry channel, allowing mRNA recruitment, and widens the P site to form the scanning-conducive py48S-open (III). eIF2β contacts eIF1 and probably stabilizes this open conformation while eIF3 undergoes major conformational change and eIF3i is repositioned on the subunit-interface. On AUG recognition, the head moves downward to clamp in the mRNA and enclose the $tRNA_i$ in the $P_{IN}$ state of py48S-closed (IV). eIF2β loses contact with eIF1 and moves away. (See text for further details.)

| **Local resolution** | **py48S-open** | **py48S-closed** | **40S•eIF1•eIF1A** | **py48S calculated from EMD-2763** (Hussain et al., 2014) |
|---|---|---|---|---|
| Overall | 6.00 | 4.90 | 3.45 | 4.00 |
| Body of 40S | 6.00 | 4.90 | 3.40 | 4.05 |
| Head of 40S | 6.30 | 5.05 | 3.55 | 4.10 |
| eIF1 | 7.90 | 5.30 | 3.65 | 6.80 |
| eIF1A | 7.20 | 5.10 | 3.70 | 4.60 |
| $tRNA_i$ | 7.90 | 5.50 | - | 5.90 |
| eIF2α | 8.75 | 6.25 | - | 6.85 |
| eIF2β | 9.15 | 7.15 | - | - |
| eIF2γ | 9.35 | 10.30 | - | 14.90 |
| ASL of $tRNA_i$ and mRNA (-4 to +4) | 7.30 | 4.80 | - | 4.70 |
| eIF3-bgi subcomplex | 10.30 | 8.30 | - | - |
| eIF3c 5-helix bundle | 7.05 | 5.70 | - | - |
| eIF3 PCI domains | - | 13.80 | - | - |

**Table 6.1.1. Local resolution, Related to Figure 2.10.1.**

## 6.2 SUPPLEMENTARY MOVIE LEGENDS

**Movie S1: py48S-closed complex, related to Figure 2.10.1**

This movie shows a 360° rotation of the map of the py48S-closed complex, followed by the fitting of refined coordinates in the map. The fitting of ligands can also be observed with the maps shown as a transparent surface. The β-propeller of eIF3b shown at the solvent interface is modeled based on previously reported structure (Aylett et al., 2015).

**Movie S2: py48S-open complex, related to Figure 2.10.1**

This movie shows a 360° rotation of the map of py48S-open complex, followed by the fitting of refined coordinates in the map. The fitting of ligands can also be observed with maps shown as a transparent surface.

**Movie S3: Morphing of PICs: py48S-open to py48S-closed complex, related to Figure 2.10.2** This movie shows the morphing of 18S rRNA in the py48S-open to the py48S-closed complex (colored cyan in the first frame). A short region (1148-1163; 1615-1627) in helix h28 is shown in red. Most ligands (except $tRNA_i$ and eIF2α) and all ribosomal proteins have been removed for clarity. The front view shows the upward movement of the head while no major conformational change is observed in the body. The change in position of $tRNA_i$ and eIF2α with the head movement can be clearly seen.

**Movie S4: Morphing of PICs: 40S eIF1 eIF1A PIC to py48S-closed complex, related to Figure 2.10.2** This movie shows the morphing of 18S rRNA in the 40S eIF1 eIF1A PIC to the py48S-closed complex (colored blue in the first frame). A short region (1148-1163; 1615-1627) in helix h28 is shown in red. All ligands and ribosomal proteins have

been removed for clarity. The front view shows the rotation of the head while no major conformational change is observed in the body.

**Movie S5: Morphing of ligands: py48S-open to py48S-closed complex, related to Figure 2.10.5**

This movie shows the morphing of the 18S rRNA and ligands in the py48S-open to the py48S-closed complex. The ligands are shown in color, as in Figure 2.10.1. Only the eIF3i subunit is shown for eIF3. All ribosomal proteins have been removed for clarity. This movie shows the conformational change that TC undergoes during the transition from the open to the closed state.

Supplementary movies can be accessed at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4534855/#mmc1.

## 6.3 SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### 6.3.1 Recombinant eIF3 production

In order to obtain an initiation complex in an open scanning-competent state, eIF5 was omitted in the preparation. Since eIF5 often copurifies with eIF3 in *S. cerevisiae* (Acker et al., 2007; Hussain et al., 2014), we overexpressed eIF3 in *Escherichia coli* as follows. The genes for subunits of eIF3 were cloned into two different but compatible polycistronic vectors: eIF3a and eIF3c were cloned into a pCDF Duet vector including an N-terminal his-tag for subunit a, and eIF3b, eIF3g and eIF3i in a pQlink vector. These two plasmids were used to transform *E. coli* Rosetta cells and the expression was carried out at 30 °C after induction with 0.5 mM IPTG. The protein was purified using the same

steps used for the protein expressed in yeast (Mitchell et al., 2010). The protein obtained is soluble, pure and seems to be expressed in stoichiometric amounts (judged by SDS-PAGE). However visualization of the purified protein by SDS-PAGE shows that the eIF3c and the eIF3g subunits are slightly smaller than expected. Mass spectrometry analysis of these bands suggests that they may be missing part of their N-terminal region due to possible proteolysis (up to 42 N-terminal residues of its 812 amino acids for eIF3c and up to 65 N- terminal residues of its 274 amino acids for eIF3g). Nonetheless, the protein seems to be functional: it is able to interact with the other eIF3 subunit (subunit eIF3j), with the 40S ribosomal subunit, and it promotes mRNA recruitment using the assay described by (Mitchell et al., 2010), with a $K_{extent}$ (defined as the concentration of eIF3 necessary for half maximal extent of mRNA recruitment) similar to that of eIF3 purified from yeast.

### 6.3.2 Reconstitution of 48S complexes

*K. lactis* 40S subunits were prepared as described earlier (Fernandez et al., 2014). *S. cerevisiae* eIF3 and eIF2 were expressed in yeast while eIF1, eIF1A and eIF5 were expressed in *E. coli* as recombinant proteins and purified as described (Acker et al., 2007). Recombinant eIF3 used for preparation of py48S-open complex was expressed in *E. coli.* Wild type $tRNA_i$ was expressed and purified from yeast and mutant $tRNA_i$ was transcribed and aminoacylated as described (Acker et al., 2007). Unstructured mRNAs with AUG (5 GGAA[UC]4UAUG[CU]4C 3 ) and AUC

(5 GGAA[UC]4UAUC[CU]4C 3 ) codons were commercially synthesized by Integrated DNA Technologies and used for the py48S-closed and py48S-open complexes,

respectively. Both complexes were reconstituted by incubating 120 nM 40S with eIF1, eIF1A, TC (consisting of eIF2, GDPCP and Met-tRNA$_i$), eIF3, eIF5 and mRNA in 40S:eIF1:eIF1A:TC:eIF3:eIF5:mRNA molar ratios of 1:2.5:2.5:1.5:1.2:2.5:2, with the exception that eIF5 was excluded in py48S-open, in 20 mM MES, pH 6.5, 40 mM potassium acetate, 10 mM ammonium acetate, 8mM magnesium acetate, 2mM dithiothreitol. The sample was used directly to make cryo-EM grids without further purification. Grids with sample for electron microscopy were prepared as described (Fernandez et al., 2014).

### 6.3.3 Analysis, structure determination, model building and refinement details of 3D classification

For py48S-closed complex data set

From about 5500 micrographs, a total of approximately 1,200,000 particles were picked. 2D class averaging was performed and aberrant particles were discarded. An initial reconstruction was made from all selected particles (1,182,309) after 2D class averaging using the yeast 40S crystal structure low pass filtered to 40 Å as an initial model. Next, a 3D classification into 8 classes with fine angular sampling was performed. Upon refinement only two classes were refined to high resolution: class 3 [12.6 %; 149,369 particles; 4.6 Å; PIC with TC] and class 4 [31.6 %; 374,737 particles; similar to PIC-2 (Hussain et al., 2014)]. The class 3, which showed a PIC with TC, was further classified into 3 classes: 3A [15,044 particles; 7.4 Å], 3B [112,924 particles; 4.3 Å; similar to py48S (Hussain et al., 2014)] and 3C [21,401 particles; 4.9 Å; py48S-closed complex].

For py48S-open complex data set

A data set of more than 2000 images was collected and about 500,000 particles were picked. An initial reconstruction was made from all selected particles (460,079) after 2D class averaging using the yeast 40S crystal structure low pass filtered to 40 Å as an initial model. Next, a 3D classification into 10 classes with fine angular sampling was performed. Class 1 showed the presence of a 40S dimer and was discarded. Class 2 showed the presence of TC. Classes 3-10 were not homogenous enough and showed the presence of at least eIF1A. Class 2 was subsequently divided into 5 classes: 2A, 2B, 2C, 2D and 2E. Class 2A was comprised of 97,864 particles and showed the presence of TC. It was again further divided into 3 classes: 2A-1, 2A-2 and 2A-3. Class 2A-1, comprised of 86,055 particles, consisted of the 40S eIF1 eIF1A complex and was refined to a resolution of 3.5 Å. Class 2A-2 was comprised of 5,174 particles and shows a PIC containing TC but without eIF3, similar to py48S (Hussain et al., 2014), at about 7 Å resolution. Surprisingly this class does not seem to have density for eIF1. Class 2A-3 was not homogenous enough to be refined to moderate resolution. Hence we made a subset of 351,827 particles by combining Classes 3-8 from the first round of classification, Classes 2B-2E from the second round and Class 2A-3 from the third round. In other words, we left out Class 1 (which contained 40S dimers); Class 2A-1 (40S eIF1 eIF1A) and Class 2A-2 (PIC with eIF1A and TC but without eIF3 and eIF1). This subset was then divided into 5 classes: A, B, C, D and E. Class A contained TC (70,365 particles) and it was then subsequently classified into 4 classes: A1, A2, A3 and A4. Only class A1 (6,127 particles; 6.1 Å) showed the presence of TC with eIF3. Class A1 was then further classified into 2 classes: A1-1 and A1-2. Class A1-1 (4,547 particles; 6.0 Å) represented

the most complete class in this data set and contained 40S with eIF1, eIF1A, TC and eIF3, described here as the py48S-open complex, while Class A1-2 (1,580 particles; 14.9 Å) contains a PIC with eIF1, eIF1A and TC (without eIF3) corresponding to the previously reported pm43S complex (Hashem et al., 2013).

### 6.3.4 Detailed model building

Initially, the atomic coordinates of py48S (PDB: 3J81) were placed into the EM density of py48S-closed complex by rigid-body fitting using Chimera (Pettersen et al., 2004). For py48S-open, the body and head of the 40S of this same model were independently placed. Previously, 40S model had been further improved using the 40S eIF1 eIF1A structure at 3.5 Å presented here. Then, each chain of the model (including ribosomal proteins, rRNA segments, protein factors and $tRNA_i$ and mRNA) was rigid-body fitted in Coot (Emsley et al., 2010) to overcome local differences in its positions. When necessary, also each separate domain of proteins was also subject to independent rigid body fitting, as was the case of factor eIF2α.

Most of eIF2β was not present in py48S (PDB: 3J81). In both open and closed complexes the relative orientation of eIF2β and eIF2γ is the same. There are three different published archaeal IF2 βγ (aIF2βγ) dimer crystal structures (Stolboushkina et al., 2008; Sokabe et al., 2006; Yatime et al., 2007), which substantially differ in the relative orientation of the two subunits. The most similar to ours correspond to PDB: 3CW2 (Stolboushkina et al., 2008), hence we superimposed aIF2βγ using our eIF2γ as a reference, and then we rigid-body fit the β-subunit independently. However, in this crystal structure used as a model, the ZBD is disordered. Therefore we used the ZBD in

PDB: 2D74 to model it.

In py48S-open, wild type $tRNA_i$ was used from PDB: 1YFG for initial rigid-body fitting into its corresponding density and further improvement of the fitting was done with the morphing tool in Coot. Also in py48S-open NTT of eIF1A was removed from the model and eIF1 model was substituted by its counterpart in the 40S eIF1 eIF1A structure. Finally, we observed a density close to the bases U and A from the anticodon of the $tRNA_i$. We reason this density most likely belongs to bases A and U from the mRNA and in consequence this fragment of mRNA was included in the final py48S-open model.

### 6.3.5 Model Building of eIF3

Fitting of eIF3a/eIF3c PCI dimer

Although the overall densities for eIF3 at the subunit interface were similar in the two structures, the higher-resolution py48S-closed map at 4.9 Å was used to generate a model for eIF3 bound in these initiation complexes. First the dimer of complete PCI domains of eIF3a and eIF3c was generated using the crystal structures of the full PCI domain of eIF3a (PDB: 4U1D) and the eIF3a/eIF3c PCI dimer (PDB: 4U1C) from *S. cerevisiae*. (Erzberger et al., 2014). This complete eIF3a/eIF3c PCI dimer was docked as a rigid body into the density on the solvent face of the 40S in py48S-close complex. Because of variation in the resolution of the eIF3 domains, we cannot resolve individual helices for the eIF3a/eIF3c PCI domains, however, it is possible to discern the overall shape and dimensions of PCI domains and do a rigid-body fit of the eIF3a/eIF3c PCI dimer. The fitting is similar to that of the eIF3a/eIF3c PCI heterodimers in the recent yeast 40S eIF1 eIF1A eIF3 structure (Aylett et al., 2015) and also to the PCI MPN core

model (PDB: 3J7K) docked into the pm43S EM map (Hashem et al., 2013).

Fitting of eIF3i and associated eIF3b-CTD and eIF3g-NTD

Density for a β-propeller domain was observed near h44 in the vicinity of eIF2γ. There are β-propeller domains in two subunits of eIF3: eIF3b and eIF3i. Crystal structures of both domains (PDB: 4U1E, 4U1F) from *S. cerevisiae* are now available (Erzberger et al., 2014). The nine-bladed β-propeller domain of eIF3b has been well documented to interact at the solvent (Liu et al., 2014; Erzberger et al., 2014) rather than the intersubunit face of the 40S as also observed in the yeast 40S eIF1 eIF1A eIF3 structure (Aylett et al., 2015). Moreover, the nine-bladed β- propeller domain of eIF3b is larger than the observed density near h44. In fact, rigid body fitting of the β-propeller domain of eIF3b shows a steric clash with 40S.

In contrast, the seven-bladed β-propeller domain of eIF3i fits well into the density. At the local resolution, it is not possible to discern its individual β-strands but the overall shape and dimensions guides the fitting of eIF3i. In one of the two crystal structures of eIF3i (PDB: 3ZWL; (Herrmannova et al., 2012)), there is a loop (residues 258- 273) emanating from the β-propeller. A single mutation of a residue of this eIF3i loop confers a severe decrease of translation initiation without affecting the integrity of eIF3 (Cuchalova et al., 2010). We observe density for this loop in a slightly different conformation and interacting with h44. eIF3i also makes an interaction with a long helix at the C-terminus of eIF3b (PDB: 4U1E, 3ZWL). We observe density for

this helix, interacting with the β-propeller domain further supporting that this density belongs to eIF3i and not eIF3b. Extra density was also observed for a portion of eIF3g-

NTD in direct contact with this β-propeller domain that is consistent with the eIF3b-CTD/eIF3i/eIF3g-NTD trimeric crystal structure (PDB: 4U1E). These observations strongly suggest that the β-propeller domain is part of eIF3i. Moreover, its contacts with eIF3b and eIF3g make it possible to orient the β-propeller domain despite the fact that the individual blades cannot be resolved.

The C-terminal helix of this stretch of eIF3g-NTD points towards the entry channel, where the remainder of the protein (not resolved here) likely binds. Similarly, the N- terminal end of the eIF3b-CTD segment in the trimeric subcomplex points towards the likely position of the eIF3b β-propeller domain and these further supports the positioning of eIF3b-CTD/eIF3i/eIF3g-NTD trimer at the subunit interface near h44 in the vicinity of eIF2γ. As we observe density for only a portion of the eIF3bβ- propeller domain in py48S-close complex, the exact length of the connector between the two eIF3b domains (eIF3b β-propeller domain and eIF3b-CTD segment) resolved here cannot be specified. However, 39 residues in this connector are adequate to span the distance as an unstructured linker.

Fitting of helices of eIF3a and eIF3c at the intersubunit interface

The eIF3a/eIF3c PCI dimer and eIF3b-CTD/eIF3i/eIF3g-NTD trimer make up more than half of the eIF3 complex in *S. cerevisiae*. Apart from the density corresponding to the eIF3b-CTD/eIF3i/eIF3g-NTD trimer, we observe three additional and differentiated regions of densities at the intersubunit interface of the 40S. One of these corresponds to a group of 5 helices that is clearly recognizable at this resolution near h21/h24/h27. More density is located in contact eIF1 on the platform and seems to

correspond to a globular domain of around 70-100 residues. Finally, there is density for two very long helices (clearly recognizable at this resolution) arranged as a coiled coil spanning the β-propeller of eIF3i and the density on the platform near eIF1.

Having assigned both the eIF3a/eIF3c PCI dimer and the eIF3b-CTD/eIF3i/eIF3g- NTD trimer, the C-terminus of eIF3a (496-964), NTD and β-propeller domain of eIF3b, N-terminus of eIF3c (1-250) and eIF3g-CTD are not accounted for. eIF3b can be ruled out because both the NTD and β-propeller domains of eIF3b are expected to be at the solvent interface (Liu et al., 2014; Erzberger et al., 2014). In fact, although we do not observe a distinct density for the whole β-propeller domain of eIF3b, we observe extra density at low resolution for part of the eIF2β-propeller domain in py48S-close complex at its expected position (see Movie S1). The eIF3g-CTD is known to bind near the entry channel (Cuchalova et al., 2010) and a solution structure of human eIF3g-CTD is available (PDB: 2CQ0). Based on this available structural data and its known location on the 40S, the eIF3g-CTD was also ruled out. Therefore, these unassigned densities should mainly correspond to segments of eIF3a and eIF3c.

Secondary structure prediction for the N-terminus of eIF3c (1-250) suggests a region of about 100 residues containing 5-6 helices of various lengths and at least another isolated helix close to the N-terminus, whereas the C-terminus of eIF3a (496-964) is predicted to consist of very long helices. Thus for the remaining unassigned density, we reason that the 5 helices near h21/h24/h27 belong to the 120-220 region of eIF3c based on this secondary structure prediction. This region (residues 120-220 of eIF3c) is predicted to have a group of 5 helices and the observed density clearly corresponds to a group of helices. The length of helices observed in the density also corresponds to what is

expected according to the secondary structure prediction for 120-220 region of eIF3c. We modeled individual helices into the density but were unable to determine its topology, as the connecting loops are not clear and there is no side-chain information for an unambiguously sequence assignment. However, we reason that these densities correspond to the helices in the region of residues 120-220 of eIF3c. Crosslinking data indicate eIF3c interacts with uS15 (Erzberger et al., 2014) and these helices can be easily linked to the PCI domain present on the solvent interface by a linker (~30 residues), which may interact with uS15 (Figure 2.10.6D) further supporting this assignment.

Based on the volume of density (equivalent to a globular domain of around 70-100 amino acids) in contact with eIF1 on the platform and its relative proximity to group of helices near h21/h24/h27 (tentatively assigned to region of residues120-220 of eIF3c), we suggest that it belongs to the N-terminal end of eIF3c (residues ~1-90), where it would form a direct contact with eIF1, in agreement with previous studies describing the most extreme N-terminal part of eIF3c as an interacting partner of eIF1 (Reibarkh et al., 2008; Erzberger et al., 2014). However, the density we observe is not sufficiently detailed so as to enable any model building. So, the assignment of this density in contact with eIF1 to N-terminal end of eIF3c (residues ~1-90) is primarily based on biochemical studies indicating N-terminal end of eIF3c interacts with eIF1. Secondly, its proximity to region of residues120-220 of eIF3c further supports it. Thirdly, we have ruled out more or less the rest of eIF3.

The remaining density for two long kinked helices spanning the β-propeller domain of eIF3i and the proposed eIF3c-NTD density near eIF1 can therefore only belong to eIF3a, in agreement with its secondary prediction of long stretches of helices,

most likely to its CTD (from residue 760). The density for the long helices is reasonably clear and thus we assign it to the CTD of eIF3a. In fact previous studies have suggested the existence of a spectrin domain (bundle of three long helices) at the CTD of eIF3a functioning as the docking site for the formation of the a:b:i:g subcomplex (Dong et al., 2013). A recent study suggested direct interaction of the CTD of eIF3a with eIF3i and with the NTD of eIF3c (Politis et al., 2015), in agreement with the model proposed here. These helices account for more than 100 residues, and therefore our assignment (from residue 760) places the extreme C-terminal ~100 residues of eIF3a not modeled here in the vicinity of the TC, consistent with a known eIF3a-CTD interaction with eIF2 (Valasek et al., 2002).

### 6.3.6 Model Refinement and Validation

For an optimal fitting of the models into the EM density maps we used REFMAC v5.8, which has been modified to work with EM maps in a wide range of resolutions (Amunts et al., 2014; Brown et al., 2015). For all ribosomal and protein factors, ProSMART (Nicholls et al., 2012) was used to generate idealized helical restraints and hydrogen bond restraints for β-sheets. Base pair and stacking restraints for rRNA, $tRNA_i$ and mRNAs were generated using the program LIBG (Brown et al., 2015). All restraints were maintained throughout refinement. Refinement with restraints helps to preserve the correct geometry of previously known structures as well as reduce overfitting. Therefore, in this work at the present resolutions, REFMAC essentially fixes the small clashes and geometry that occur after separate rigid body fitting of individual domains. Average Fourier shell correlation (FSC) was monitored during refinement. The result of the refinements was checked visually in Coot (Emsley et al., 2010). Final model was

validated using MolProbity (Chen et al., 2010). To prevent overfitting, the global refinement and external restraints weights were carefully adjusted by cross-validation, as previously described (Brown et al., 2015; Amunts et al., 2014). Refinement statistics are given in Table 6.1.1.

### 6.3.7 Yeast strain constructions

To generate strains LMY61, LMY74, and LMY76, strain KAY18 (*MATα leu2-3 leu2- 112 ura3-53 ino1 sui3*Δ *gcn2*Δ p921 (*SUI3*$^{+}$*, URA3*)) (Asano et al., 1999) was transformed to Leu$^{+}$ with low-copy (lc) *LEU2* plasmids harboring FLAG-tagged alleles *SUI3-FL* (plasmid YCpSUI3), *sui3-FL-S202A,K214A* (pLEM13), or *sui3-FL-F217A,Q221A* (pLEM15), respectively, and the resident *SUI3*$^{+}$*,URA3* plasmid p921 was evicted by selection on 5-fluoorotic acid (5-FOA) medium. To generate strains ATY49, ATY53, and ATY54, strain JCY03 (*MATa ura3-52 leu2-3 leu2-112 trp1*Δ*- 63 his4-301(ACG) sui1*Δ*::hisG* p1200 (*SUI1*$^{+}$*, URA3*)) (Cheung et al., 2007) was transformed to Leu$^{+}$ with single-copy *LEU2* plasmids harboring *SUI1*$^{+}$ (pJCB101), *sui1-F108A* (pAT117), or *sui1-F108D* (pAT118), respectively, and the resident *SUI1*$^{+}$ *URA3* plasmid (p1200) was evicted by selecting for growth on 5-FOA medium. Plasmids pLEM13 and pLEM15 were constructed from YCpSUI3 (Asano et al., 1999) using the QuikChange site-directed mutagenesis system (Stratagene) according to the manufacturer's directions and the appropriate primers. Plasmids pAT117 and pAT118 were similarly constructed from pJCB101 (Martin-Marcos et al., 2011).

### 6.3.8 Yeast biochemical methods

Assays of β-galactosidase activities in whole-cell extracts (WCEs) were

performed as described previously (Moehle and Hinnebusch, 1991) using transformants harboring the appropriate reporter plasmids, *HIS4$_{AUG}$-lacZ* (p367, *HIS4$_{UUG}$-lacZ* (p391) (Donahue and Cigan, 1988), or *GCN4-lacZ* (p180) (Hinnebusch, 1985). Transformants were cultured in appropriately supplemented synthetic dextrose minimal media (SD) at 30°C to an $A_{600}$ of ~0.8. The same culture conditions were used for Western or coimmunoprecipitation analyses. WCEs for Western analysis were prepared by trichloroacetic acid extraction as previously described (Reid and Schatz, 1982), and immunoblot analysis was conducted as described (Nanda et al., 2009) using antibodies against Flag epitope (Sigma), eIF2Bε (Bushman et al., 1993), eIF1 (Valasek et al., 2004), or eIF3j (Valasek et al., 2001). Coimmunoprecipitations were conducted as previously described (Asano et al., 1999) and immunoblot analysis of immune complexes was conducted as above using antibodies against Flag epitope, eIF2γ (Hannig et al., 1993), eIF2α (Dever et al., 1995), and eIF2Bε.

## 6.4 SUPPLEMENTAL REFERENCES

Acker, M. G., Kolitz, S. E., Mitchell, S. F., Nanda, J. S., and Lorsch, J. R. (2007). Reconstitution of yeast translation initiation. Methods Enzymol *430*, 111-145.

Amunts, A., Brown, A., Bai, X. C., Llacer, J. L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S. H., and Ramakrishnan, V. (2014). Structure of the yeast mitochondrial large ribosomal subunit. Science *343*, 1485-1489.

Asano, K., Krishnamoorthy, T., Phan, L., Pavitt, G. D., and Hinnebusch, A. G. (1999). Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. EMBO J *18*, 1673-1688.

Aylett, C. H., Boehringer, D., Erzberger, J. P., Schaefer, T., and Ban, N. (2015). Structure of a Yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. Nat Struct Mol Biol

Brown, A., Long, F., Nicholls, R. A., Toots, J., Emsley, P., and Murshudov, G. (2015). Tools for macromolecular model building and refinement into electron cryo- microscopy reconstructions. Acta Crystallogr D Biol Crystallogr *71*, 136-153.

Bushman, J. L., Foiani, M., Cigan, A. M., Paddon, C. J., and Hinnebusch, A. G. (1993). Guanine nucleotide exchange factor for eukaryotic translation initiation factor 2 in Saccharomyces cerevisiae: interactions between the essential subunits GCD2, GCD6, and GCD7 and the regulatory subunit GCN3. Mol Cell Biol *13*, 4618-4631.

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr *66*, 12-21.

Cheung, Y. N., Maag, D., Mitchell, S. F., Fekete, C. A., Algire, M. A., Takacs, J. E., Shirokikh, N., Pestova, T., Lorsch, J. R., and Hinnebusch, A. G. (2007). Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. Genes Dev *21*, 1217-1230.

Cuchalova, L., Kouba, T., Herrmannova, A., Danyi, I., Chiu, W. L., and Valasek, L. (2010). The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. Mol Cell Biol *30*, 4671-4686.

Dever, T. E., Yang, W., Astrom, S., Bystrom, A. S., and Hinnebusch, A. G. (1995). Modulation of tRNA(iMet), eIF-2, and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of eIF-2.GTP.Met-tRNA(iMet) ternary complexes. Mol Cell Biol *15*, 6351-6363.

Donahue, T. F., and Cigan, A. M. (1988). Genetic selection for mutations that reduce or abolish ribosomal recognition of the HIS4 translational initiator region. Mol Cell Biol *8*, 2955-2963.

Dong, Z., Qi, J., Peng, H., Liu, J., and Zhang, J. T. (2013). Spectrin domain of eukaryotic initiation factor 3a is the docking site for formation of the a:b:i:g subcomplex. J Biol

Chem *288*, 27951-27959.

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Erzberger, J. P., Stengel, F., Pellarin, R., Zhang, S., Schaefer, T., Aylett, C. H., Cimermancic, P., Boehringer, D., Sali, A., Aebersold, R., and Ban, N. (2014). Molecular architecture of the 40SeIF1eIF3 translation initiation complex. Cell *158*, 1123-1135.

Fernandez, I. S., Bai, X. C., Murshudov, G., Scheres, S. H., and Ramakrishnan, V. (2014). Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. Cell *157*, 823-831.

Hannig, E. M., Cigan, A. M., Freeman, B. A., and Kinzy, T. G. (1993). GCD11, a negative regulator of GCN4 expression, encodes the gamma subunit of eIF-2 in Saccharomyces cerevisiae. Mol Cell Biol *13*, 506-520.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H. Y., Grassucci, R. A., Hellen, C. U., Pestova, T. V., and Frank, J. (2013). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. Cell *153*, 1108-1119.

Herrmannova, A., Daujotyte, D., Yang, J. C., Cuchalova, L., Gorrec, F., Wagner, S., Danyi, I., Lukavsky, P. J., and Valasek, L. S. (2012). Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. Nucleic Acids Res *40*, 2294-2311.

Hinnebusch, A. G. (1985). A hierarchy of trans-acting factors modulates translation of an activator of amino acid biosynthetic genes in Saccharomyces cerevisiae. Mol Cell Biol *5*, 2349-2360.

Hussain, T., Llacer, J. L., Fernandez, I. S., Munoz, A., Martin-Marcos, P., Savva, C. G., Lorsch, J. R., Hinnebusch, A. G., and Ramakrishnan, V. (2014). Structural changes enable start codon recognition by the eukaryotic translation initiation complex. Cell *159*, 597-607.

Liu, Y., Neumann, P., Kuhle, B., Monecke, T., Schell, S., Chari, A., and Ficner, R. (2014). Translation initiation factor eIF3b contains a nine-bladed beta-propeller and interacts with the 40S ribosomal subunit. Structure *22*, 923-930.

Martin-Marcos, P., Cheung, Y. N., and Hinnebusch, A. G. (2011). Functional elements in initiation factors 1, 1A, and 2beta discriminate against poor AUG context and non-AUG start codons. Mol Cell Biol *31*, 4814-4831.

Mitchell, S. F., Walker, S. E., Algire, M. A., Park, E. H., Hinnebusch, A. G., and Lorsch, J. R. (2010). The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. Mol Cell *39*, 950-962.

Moehle, C. M., and Hinnebusch, A. G. (1991). Association of RAP1 binding sites with stringent control of ribosomal protein gene transcription in Saccharomyces cerevisiae. Mol Cell Biol *11*, 2723-2735.

Nanda, J. S., Cheung, Y. N., Takacs, J. E., Martin-Marcos, P., Saini, A. K., Hinnebusch, A. G., and Lorsch, J. R. (2009). eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. J Mol Biol *394*, 268-285.

Nicholls, R. A., Long, F., and Murshudov, G. N. (2012). Low-resolution refinement tools in REFMAC5. Acta Crystallogr D Biol Crystallogr *68*, 404-417.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Politis, A., Schmidt, C., Tjioe, E., Sandercock, A. M., Lasker, K., Gordiyenko, Y., Russel, D., Sali, A., and Robinson, C. V. (2015). Topological Models of Heteromeric Protein Assemblies from Mass Spectrometry: Application to the Yeast eIF3:eIF5 Complex. Chem Biol *22*, 117-128.

Reibarkh, M., Yamamoto, Y., Singh, C. R., del Rio, F., Fahmy, A., Lee, B., Luna, R. E., Ii, M., Wagner, G., and Asano, K. (2008). Eukaryotic initiation factor (eIF) 1 carries two distinct eIF5-binding faces important for multifactor assembly and AUG selection. J Biol Chem *283*, 1094-1103.

Reid, G. A., and Schatz, G. (1982). Import of proteins into mitochondria. Extramitochondrial pools and post-translational import of mitochondrial protein precursors in vivo. J Biol Chem *257*, 13062-13067.

Sokabe, M., Yao, M., Sakai, N., Toya, S., and Tanaka, I. (2006). Structure of archaeal

translational initiation factor 2 betagamma-GDP reveals significant conformational change of the beta-subunit and switch 1 region. Proc Natl Acad Sci U S A *103*, 13016-13021.

Stolboushkina, E., Nikonov, S., Nikulin, A., Blasi, U., Manstein, D. J., Fedorov, R., Garber, M., and Nikonov, O. (2008). Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. J Mol Biol *382*, 680-691.

Valasek, L., Nielsen, K. H., and Hinnebusch, A. G. (2002). Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. EMBO J *21*, 5886-5898.

Valasek, L., Nielsen, K. H., Zhang, F., Fekete, C. A., and Hinnebusch, A. G. (2004). Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. Mol Cell Biol *24*, 9437-9455.

Valasek, L., Phan, L., Schoenfeld, L. W., Valaskova, V., and Hinnebusch, A. G. (2001). Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. EMBO J *20*, 891- 904.

Yatime, L., Mechulam, Y., Blanquet, S., and Schmitt, E. (2007). Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states. Proc Natl Acad Sci U S A *104*, 18445-18450.

# 7. Appendix B: Supplemental Material for "A Network of eIF2β Interactions with eIF1 and Met-tRNA$_i$ Promotes Accurate Start Codon Selection by the Translation Initiation Complex"

Supplemental Material for Chapter 3:

A Network of eIF2β interactions with eIF1 and Met-tRNA$_i$ promotes accurate start codon selection by the translation initiation complex

Laura Marler, Anil Thakur, Alan G. Hinnebusch

Nucleic Acids Research 2018

**SUPPLEMENTARY MATERIAL FOR:**

# A network of eIF2β interactions with eIF1 and Met-tRNA$_i$ promotes accurate start codon selection by the translation preinitiation complex

Anil Thakur, Laura Marler, and Alan G. Hinnebusch

## 7.1 SUPPLEMENTAL TABLES

**Table 7.1.1. Plasmids used in the study of eIF2β interactions with eIF1.**

| Plasmid | Description | Parent Plasmid | Source or Reference |
|---|---|---|---|
| YCplac111 | sc *LEU2* cloning vector | | (361) |
| p1200 | sc *URA3 SUI1* in YCp50 | | (209) |
| YCplac112 | hc TRP cloning Vector | | (361) |
| pJCB101 | sc *LEU2 SUI1* in YCplac111 | | (212) |
| YEp24 | hc *URA3* cloning vector | | (362) |
| ATP114 | sc *LEU2 sui1-Q31A* in YCplac111 | pJCB101 | This study |
| ATP115 | sc *LEU2 sui1-Q31K* in YCplac111 | pJCB101 | This study |
| ATP116 | sc *LEU2 sui1-Q31E* in YCplac111 | pJCB101 | This study |
| ATP117 | sc *LEU2 sui1*-F108Ain YCplac111 | pJCB101 | This study |
| ATP118 | sc *LEU2 sui1-F108D* in YCplac111 | pJCB101 | This study |
| ATP119 | sc *LEU2 sui1-F108R* in YCplac111 | pJCB101 | This study |
| ATP124 | sc *LEU2 sui1-Q31A-F108A* in YCplac111 | pJCB101 | This study |
| p367 | sc *URA3 HIS4(ATG)-lacZ* | | (360) |
| p391 | sc *URA3 HIS4(TTG)-lacZ* | | (360) |
| p180 | sc *URA3 GCN4-lacZ* in YCp50 | | (363) |
| pPMB24 | sc *URA3 SUI1-lacZ* | | (212) |
| pPMB25 | sc *URA3* $SUI1_{opt}$*-lacZ* | | (212) |
| pC3502 | sc *URA3* $^{-3}$AAA$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |
| pC4466 | sc *URA3* $^{-3}$UAA$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |
| pC3503 | sc *URA3* $^{-3}$UUU$^{-1}$ el.uORF1 *GCN4-lacZ* in YCp50 | | (314) |
| pC3505 | sc *URA3* el.uORF1-less *GCN4-lacZ* in YCp50 | | (314) |
| p1780-IMT | hc *URA3 SUI2, SUI3, GCD11, IMT4* | YEp24 | (297) |
| p4385 | hc *TRP1- SUI2, SUI3, GCD11, IMT4* | p1780-IMT | Christie Fekete |

| Plasmid | Description | Parent Plasmid | Source or Reference |
|---|---|---|---|
| p4280/YCpSUI3-S264Y-W | sc *TRP1 SUI3-S264Y* | YCplac22 | (146) |
| p4281/YCpTIF5-G31R-W | sc *TRP1 TIF5-G31R* | YCplac22 | (146) |
| pTYB2-eIF1 | *SUI1* in PTYB2 | pTYB2 | (347),(347) |
| pPMB97 | *sui1-K60E* in pTYB2 | pTYB2-eIF1 | (213) |
| ATP172 | *sui1-Q31E* in pTYB2 | pTYB2-eIF1 | This study |
| ATP173 | *sui1-F108D* in pTYB2 | pTYB2-eIF1 | This study |
| ATP174 | *sui1-F108R* in pTYB2 | pTYB2-eIF1 | This study |
| ATP175 | *sui1-Q31A-F108A* in pTYB2 | pTYB2-eIF1 | This study |
| p921 | *sc URA3 SUI3* | pRS316 | (364) |
| p920 | *sc LEU2 SUI3* | pRS315 | Thomas Dever |
| LMP24 | *sc LEU2 SUI3-E198A* | p920 | This study |
| LMP27 | *sc LEU2 SUI3-F217A* | p920 | This study |
| LMP29 | *sc LEU2 SUI3-F217A/Q221A* | p920 | This study |
| LMP32 | *sc LEU2 SUI3-K170A* | p920 | This study |
| LMP33 | *sc LEU2 SUI3-K170A/S202A* | p920 | This study |
| LMP37 | *sc LEU2 SUI3-K214A* | p920 | This study |
| LMP38 | *sc LEU2 SUI3-S202A/K214A* | p920 | This study |
| LMP94 | *sc LEU2 SUI3-E189A* | p920 | This study |
| LMP95 | *sc LEU2 SUI3-E189R* | p920 | This study |
| LMP96 | *sc LEU2 SUI3-Q193A* | p920 | This study |
| LMP97 | *sc LEU2 SUI3-Q193R* | p920 | This study |
| LMP98 | *sc LEU2 SUI3-E189A/Q193A* | p920 | This study |
| LMP99 | *sc LEU2 SUI3-E189R/Q193R* | p920 | This study |
| pAV1089 | *hc URA3(6x)His-GCD11, SUI2, SUI3* | YEp24 | (365) |
| pAV1726 | *hc LEU2 (6x)His-GCD11, SUI2, SUI3* | pRS425 | (366) |
| LMP91 | *hc LEU2 (6x)His-GCD11, SUI2, SUI3-F217A/Q221A* | pAV1726 | This study |
| LMP92 | *hc LEU2 (6x)His-GCD11, SUI2, SUI3-S202A/K214A* | pAV1726 | This study |
| LMP101 | *hc LEU2 (6x)His-GCD11, SUI2, SUI3-E189R* | pAV1726 | This study |

**Table 7.1.2. Yeast strains used in the study of eIF2β interactions with eIF1.**

| Strain | Genotype | Source |
|---|---|---|
| JCY03 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* p1200 (sc *URA3 SUI1*) | (129) |
| ATY100 (PMY30) | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* pJCB101 (sc *LEU2 SUI1*) | (212) |
| ATY122 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP114 (sc *LEU2 sui1-Q31A*) | This study |
| ATY123 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP115 (sc *LEU2 sui1-Q31K*) | This study |
| ATY124 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP116 (sc *LEU2 sui1Q31E*) | This study |
| ATY125 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP117 (sc *LEU2 sui1-F108A*) | This study |
| ATY126 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP118 (sc *LEU2 sui1-F108D*) | This study |
| ATY127 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP119 (sc *LEU2 sui1-F108R*) | This study |
| ATY128 | *MATa ura3-52 leu2-3 leu2-112 trp1Δ-63 his4-301*(*ACG*) *sui1Δ::hisG* ATP124 (sc *LEU2 sui1-Q31A-F108A)* | This study |
| KAY18 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□p921(*SUI3, URA3*) | (297) |
| LMY103 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP24(*SUI3-E198A, LEU2*) | This study |
| LMY106 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP27(*SUI3-F217A, LEU2*) | This study |
| LMY108 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP29(*SUI3-F217A/Q221A, LEU2*) | This study |
| LMY111 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP32(*SUI3-K170A, LEU2*) | This study |
| LMY112 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP33(*SUI3-K170A/S202A, LEU2*) | This study |
| LMY116 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP37(*SUI3-K214A, LEU2*) | This study |
| LMY117 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP38(*SUI3-S202A/K214A, LEU2*) | This study |
| LMY130 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP94(*SUI3-E189A, LEU2*) | This study |
| LMY131 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP95(*SUI3-E189R, LEU2*) | This study |

| Strain | Genotype | Source |
|---|---|---|
| LMY132 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP96(*SUI3-Q193A, LEU2*) | This study |
| LMY133 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP97(*SUI3-Q193R, LEU2*) | This study |
| LMY134 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP98(*SUI3-E189A/Q193A, LEU2*) | This study |
| LMY135 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□LMP99(*SUI3-E189R/Q193R, LEU2*) | This study |
| LMY142 | *MATα leu2-3 leu2-112 ura3-53 ino1 sui3Δ gcn2Δ*□p920(*SUI3, LEU2*) | This study |
| H3840 | *MATα leu2-3 leu2-112 ura3-52 ino1 sui2Δ gcn2Δ pep4::leu2::NatMX4 sui3::KANMX4 <HIS4-lacZ, ura3-52>* pAV1089(*SUI2, SUI3, [6x]-GCD11, URA3)* | (214) |
| LMY128 | *MATα leu2-3 leu2-112 ura3-52 ino1 sui2Δ gcn2Δ pep4::leu2::NatMX4 sui3::KANMX4 <HIS4-lacZ, ura3-52>* LMP92(*SUI2, SUI3-S202A/K214A, [6x]-GCD11, URA3)* | This study |
| LMY129 | *MATα leu2-3 leu2-112 ura3-52 ino1 sui2Δ gcn2Δ pep4::leu2::NatMX4 sui3::KANMX4 <HIS4-lacZ, ura3-52>* LMP91(*SUI2, SUI3-F217A/Q221A, [6x]-GCD11, URA3)* | This study |
| LMY137 | *MATα leu2-3 leu2-112 ura3-52 ino1 sui2Δ gcn2Δ pep4::leu2::NatMX4 sui3::KANMX4 <HIS4-lacZ, ura3-52>* LMP101(*SUI2, SUI3-E189R, [6x]-GCD11, URA3)* | This study |

**Table 7.1.3. Oligonucleotide primers used in the study of eIF2β interactions with eIF1.**

| SUBSTITUTION | SEQUENCE 5' -3' |
|---|---|
| **eIF1-Q31A FOR*** | CTATATTCATATTCGTATCGCACAGAGAAATGGTAGA AAAACTTTAACTAC |
| **eIF1-Q31A REV*** | GTAGTTAAAGTTTTTCTACCATTTCTCTGTGCGATACG AATATGAATATAG |
| **eIF1-Q31KFOR** | CTATATTCATATTCGTATCAAACAGAGAAATGGTAGA AAAACTTTAACTACGG |
| **eIF1-Q31K REV** | CCGTAGTTAAAGTTTTTCTACCATTTCTCTGTTTGATAC GAATATGAATATAG |
| **eIF1-Q31E FOR** | CAAACTATATTCATATTCGTATCGAACAGAGAAATGG TAGAAAAACTTTAAC |
| **eIF1-Q31E REV** | GTTAAAGTTTTTCTACCATTTCTCTGTTCGATACGAAT ATGAATATAGTTTG |
| **eIF1-F108A FOR** | GAAGAACATTAAAATTCATGGGGCTTAAGTTCAAGGC TTACGCCGAGTC |
| **eIF1-F108A REV** | GACTCGGCGTAAGCCTTGAACTTAAGCCCCATGAATTT TAATGTTCTTC |
| **eIF1-F108D FOR** | GAAGAACATTAAAATTCATGGGGATTAAGTTCAAGGC TTACGCCGAGTC |
| **eIF1-F108D REV** | GACTCGGCGTAAGCCTTGAACTTAATCCCCATGAATTT TAATGTTCTTC |
| **eIF1-F108R FOR** | GAAGAACATTAAAATTCATGGGCGTTAAGTTCAAGGC TTACGCCGAGTC |
| **eIF1-F108R REV** | GACTCGGCGTAAGCCTTGAACTTAACGCCCATGAATTT TAATGTTCTTC |
| **eIF2β-E198A FOR** | TTCAATATCTCTTCGCAGCATTAGGTACGTCCGGTTC |
| **eIF2β-E198A REV** | GAACCGGACGTACCTAATGCTGCGAAGAGATATTGAA |
| **eIF2β-F217A FOR** | GAAAAGATTAGTCATTAAGGGTAAGGCTCAATCCAAA CAAATGGAGAATGTC |
| **eIF2β-F217A REV** | GACATTCTCCATTTGTTTGGATTGAGCCTTACCCTTAA TGACTAATCTTTTC |
| **eIF2β-F217A/Q221A FOR** | GTCAGAAAAGATTAGTCATTAAGGGTAAGGCTCAATC CAAAGCAATGGAGAATGTCTTAAGAAGATACATTT |

| SUBSTITUTION | SEQUENCE 5’ -3’ |
|---|---|
| **eIF2β-F217A/Q221A REV** | AAATGTATCTTCTTAAGACATTCTCCATTGCTTTGGATTGAGCCTTACCCTTAATGACTAATCTTTTCTGAC |
| **eIF2β-K170A FOR** | CTCCTGTTTGTTTGCGTGATGGTGCGAAGACTATTTTCTCGAATATCC |
| **eIF2β-K170A REV** | GGATATTCGAGAAAATAGTCTTCGCACCATCACGCAAACAAACAGGAG |
| **eIF2β-S202A FOR** | CGCAGAATTAGGTACGGCCGGTTCTGTTGACGG |
| **eIF2β-S202A REV** | CCGTCAACAGAACCGGCCGTACCTAATTCTGCG |
| **eIF2β-K214A FOR** | TTCTGTTGACGGTCAGAAAAGATTAGTCATTGCGGGTAAGTTTCAATCC |
| **eIF2β-K214 REV** | GGATTGAAACTTACCCGCAATGACTAATCTTTTCTGACCGTCAACAGAA |
| **eIF2β-E189A FOR** | CGAAAAATTGCATAGATCTCCGGCACATTTGATTCAATATCTCTTCG |
| **eIF2β-E189A REV** | CGAAGAGATATTGAATCAAATGTGCCGGAGATCTATGCAATTTTTCG |
| **eIF2β-E189R FOR** | GCCGAAAAATTGCATAGATCTCCGAGACATTTGATTCAATATCTCTTCGC |
| **eIF2β-E189R REV** | GCGAAGAGATATTGAATCAAATGTCTCGGAGATCTATGCAATTTTTCGGC |
| **eIF2β-Q193A FOR** | CATAGATCTCCGGAACATTTGATTGCATATCTCTTCGCAGAATTAGGTA |
| **eIF2β-Q193A REV** | TACCTAATTCTGCGAAGAGATATGCAATCAAATGTTCCGGAGATCTATG |
| **eIF2β-Q193R FOR** | GATCTCCGGAACATTTGATTCGATATCTCTTCGCAGAATTAGG |
| **eIF2β-Q193R REV** | CCTAATTCTGCGAAGAGATATCGAATCAAATGTTCCGGAGATC |
| **eIF2β-E189A/Q193A FOR** | CGAAAAATTGCATAGATCTCCGGCACATTTGATTGCATATCTCTTCGCAGAATTAGGTA |

| SUBSTITUTION | SEQUENCE 5’ -3’ |
|---|---|
| **eIF2β-E189A/Q193A REV** | TACCTAATTCTGCGAAGAGATATGCAATCAAATGTGCCGGAGATCTATGCAATTTTTCG |
| **eIF2β-E189R/Q193R FOR** | CGAAAAATTGCATAGATCTCCGAGACATTTGATTCGATATCTCTTCGCAGAATTAGGTA |
| **eIF2β-E189R/Q193R REV** | TACCTAATTCTGCGAAGAGATATCGAATCAAATGTCTCGGAGATCTATGCAATTTTTCG |

*FOR, forward primer ; REV, reverse primer

## 7.2 SUPPLEMENTAL FIGURES

Fig. S1

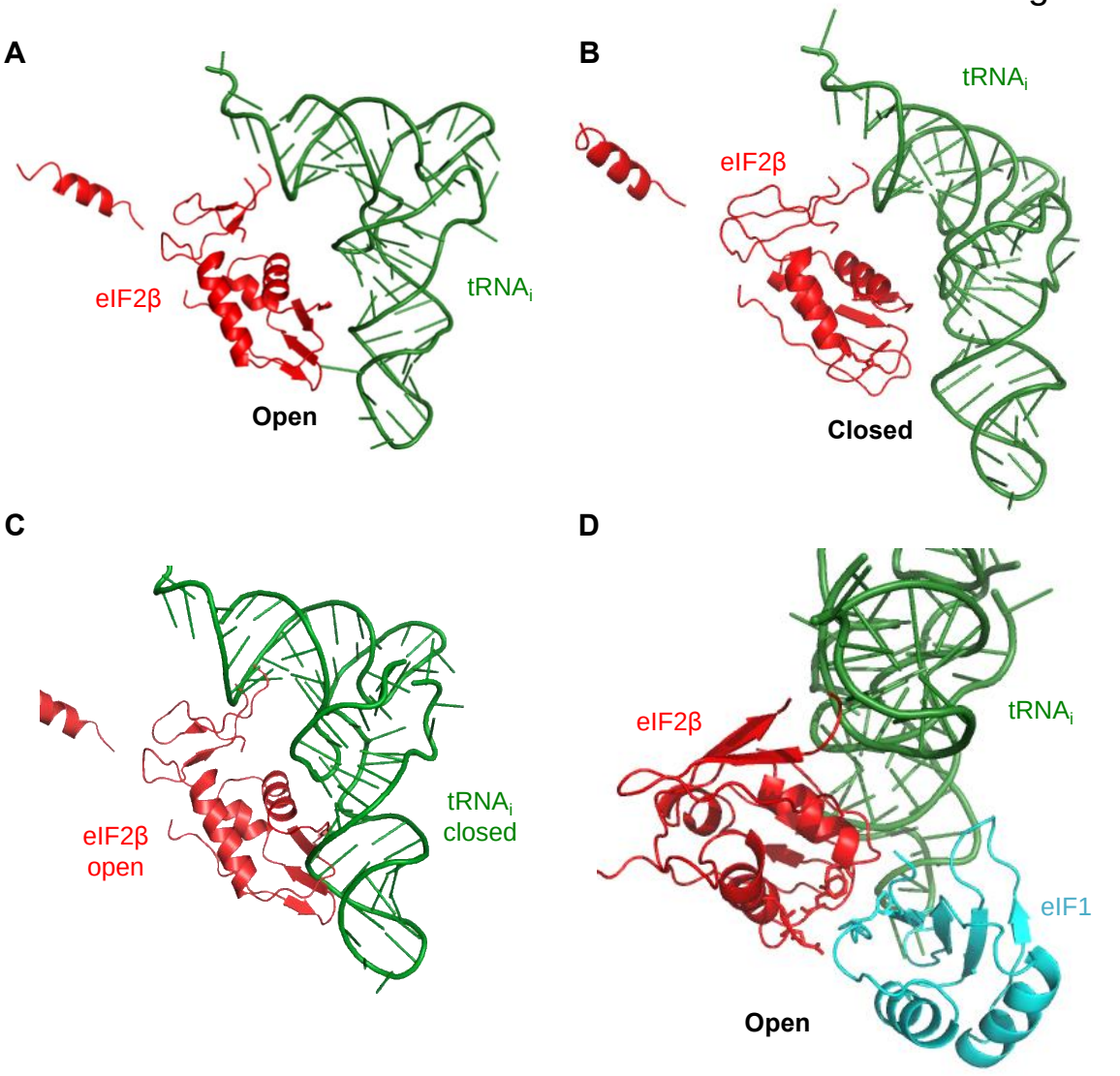


**Figure 7.2.1. Distinct interactions of eIF2β with $tRNA_i$ in py48S-open versus py48S-closed should enable eIF2β to restrict transition to the closed/$P_{IN}$ conformation of the PIC in a manner facilitated by eIF1.**

**(A)** Interactions between eIF2β and the $tRNA_i$ ASL and D-loop in the py48S-open complex. **(B)** Interactions of eIF2β with $tRNA_i$ in py48S-closed differ from those in (A) and are restricted to the D-loop. **(C)** An overlay of eIF2β in py48S-open with $tRNA_i$ in py48S-closed reveals predicted clashes throughout the ASL. These clashes are alleviated during the open-to-closed transition by movement of eIF2β both laterally and toward the D-loop. The predicted clashes suggest that eIF2β performs a steric role in ensuring acccurate start codon selection, undergoing a conformational change that allows $tRNA_i$ to assume the $P_{IN}$ state of the closed conformation only when a perfect AUG:anticodon duplex is formed in the P-site. **(D)** A network of interactions between eIF2β and both eIF1 and $tRNA_i$ in py48S-open enhances TC recruitment and stabilizes the open, scanning conformation of the PIC prior to AUG selection.

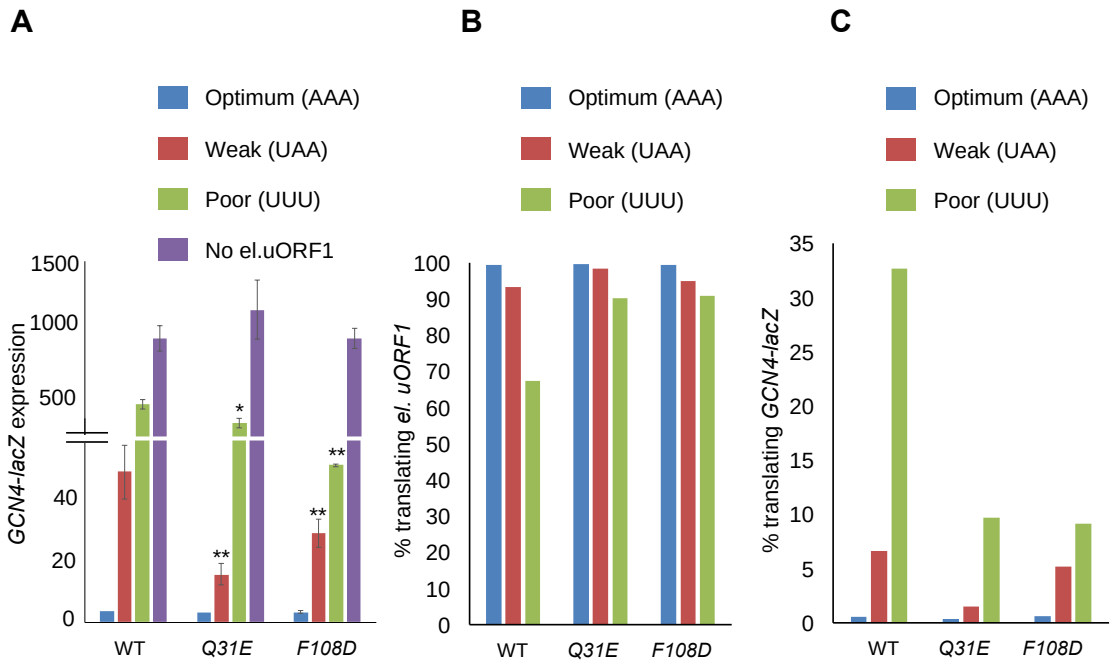


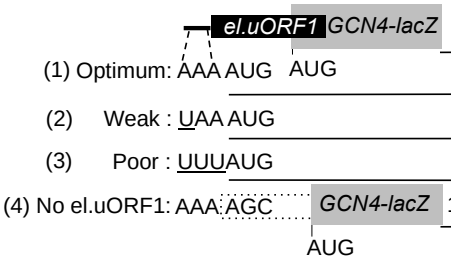

**Figure 7.2.2. eIF1 substitutions Q31E and F108D decrease discrimination against the *GCN4* uORF1 AUG codon in suboptimal context.**

**(A)** Transformants of JCY03 harboring WT *SUI1*, *sui1-Q31E* or *sui1-F108D* and el.uORF1 *GCN4-lacZ* reporters (pC3502, pC3503 or pC4466) containing, respectively, optimum, weak or poor context of uAUG-1, or an uORF-less *GCN4-lacZ* reporter with a mutated uAUG-1 (pC3505), were assayed for β-galactosidase activities as in Fig. 3.10.2 D. Mean expression values with SEMs were determined from six transformants and asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t test (*$P < 0.05$; **$P < 0.01$). **(B, C)** The percentages of

scanning ribosomes that translate el.uORF1 (B) or leaky-scan uAUG-1 and translate *GCN4-lacZ* (C) were calculated from the data in (A) by comparing the amount of *GCN4-lacZ* expression observed for each uORF-containing reporter to the uORF-less construct, yielding the percentages in (C), and subtracting the values in (C) from 100 to obtain the percentages in (B).

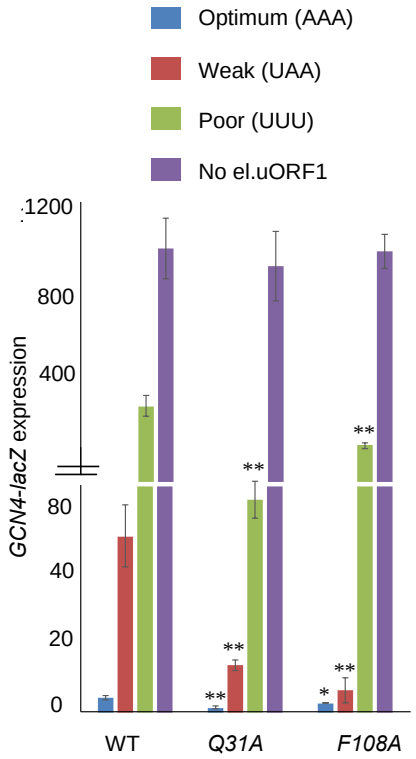
Fig. S3

**A**

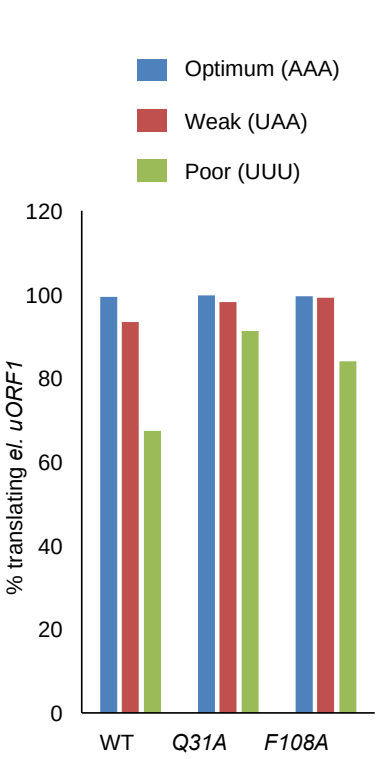



| | *GCN4-lacZ* expression(U) | | | % translating *GCN4-lacZ* | | | % translating *el. uORF1* | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) WT | (2) *Q31A* | (3) *F108A* | (4) WT | (5) *Q31A* | (6) *F108A* | (7) WT | (8) *Q31A* | (9) *F108A* |
| (1) Optimum: AAA AUG AUG | 5±0.7 | 2±0.6** | 3±0.2* | 0.5 | 0.2 | 0.3 | 99.5 | 99.8 | 99.7 |
| (2) Weak : UAA AUG | 66±11 | 17±2** | 8±5** | 8 | 2 | 1 | 92 | 98 | 99 |
| (3) Poor : UUUAUG | 327±43 | 80±7** | 158±15** | 33 | 9 | 16 | 67 | 91 | 84 |
| (4) No el.uORF1: AAA AGC GCN4-lacZ AUG | 1000±150 | 900±200 | 1000±100 | >99 | >99 | >99 | <1 | <1 | <1 |

**B**




**C**




**D**

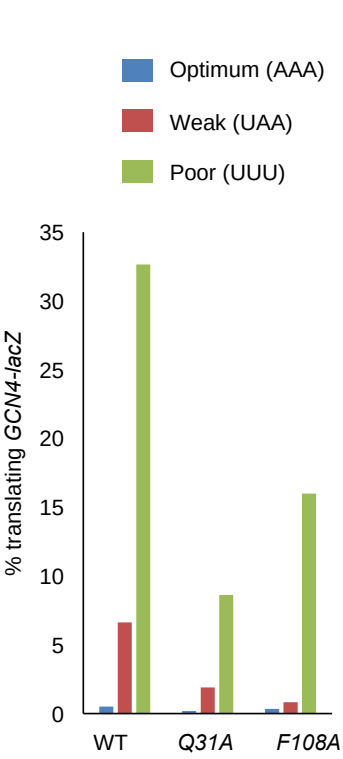


**Figure 7.2.3. eIF1 substitutions Q31A and F108A decrease discrimination against the *GCN4* uORF1 AUG codon in suboptimal context.**

**(A)** Transformants of JCY03 harboring WT *SUI1*, *sui1-Q31A* and *eIF1-F108A* and the el. uORF1 *GCN4-lacZ* reporters containing optimum (row 1), weak (row 2) or poor (row 3) context of uAUG-1, or an uORF-less *GCN4-lacZ* reporter with a mutated uAUG-1 (row 4), were analyzed as in Fig. 3.10.3 C. **(B)** β-galactosidase activities from columns 1-3 of (A) plotted in graphical format for WT *SUI1*, *sui1*-Q31A or *sui1-F108A* transformants containing the el. uORF1 *GCN4-lacZ* reporters with optimum, weak or poor context of uAUG-1, or an uORF-less *GCN4-lacZ* reporter with a mutated uAUG. Mean expression values with SEMs were determined from six transformants and asterisks indicate significant differences between mutant and WT as judged by a two-tailed, unpaired Student's t test (*P < 0.05; **P <0.01). **(C, D)** The percentages of scanning ribosomes that translate el.uORF1 (C) or leaky-scan uAUG-1 and translate *GCN4-lacZ* (D) were calculated from the data in (B) as described in Fig. 7.2.2 B-C.

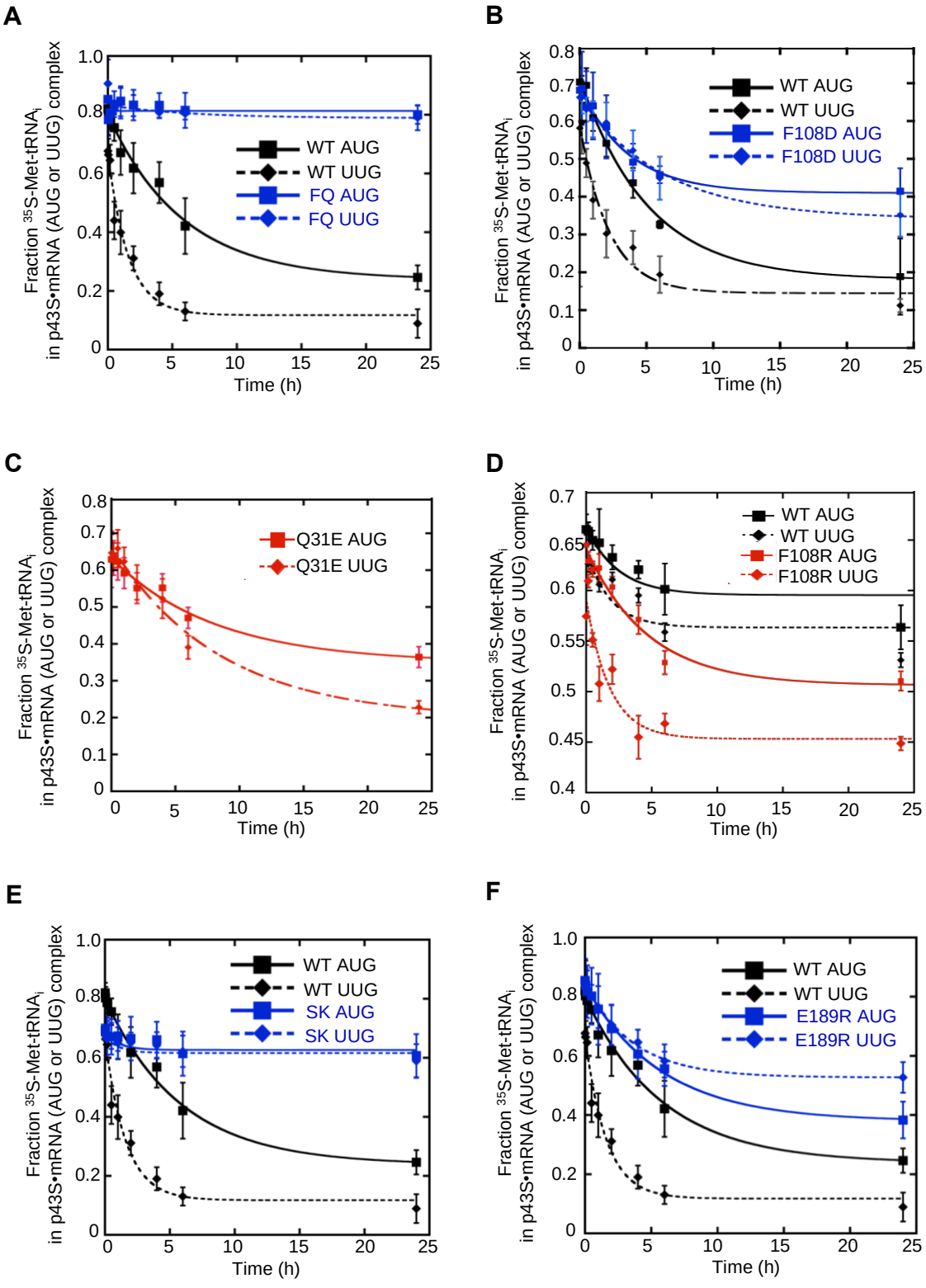
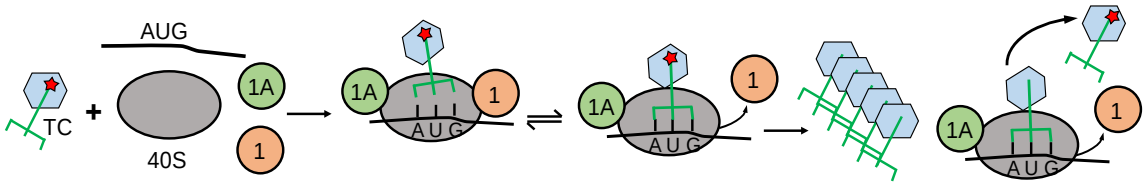

A
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
WT AUG
WT UUG
FQ AUG
FQ UUG
B
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
WT AUG
WT UUG
F108D AUG
F108D UUG
C
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
Q31E AUG
Q31E UUG
D
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
WT AUG
WT UUG
F108R AUG
F108R UUG
E
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
WT AUG
WT UUG
SK AUG
SK UUG
F
Fraction $^{35}$S-Met-tRNA$_i$ in p43S•mRNA (AUG or UUG) complex
Time (h)
WT AUG
WT UUG
E189R AUG
E189R UUG

**Figure 7.2.4. Plots of TC dissociation assays with error.**

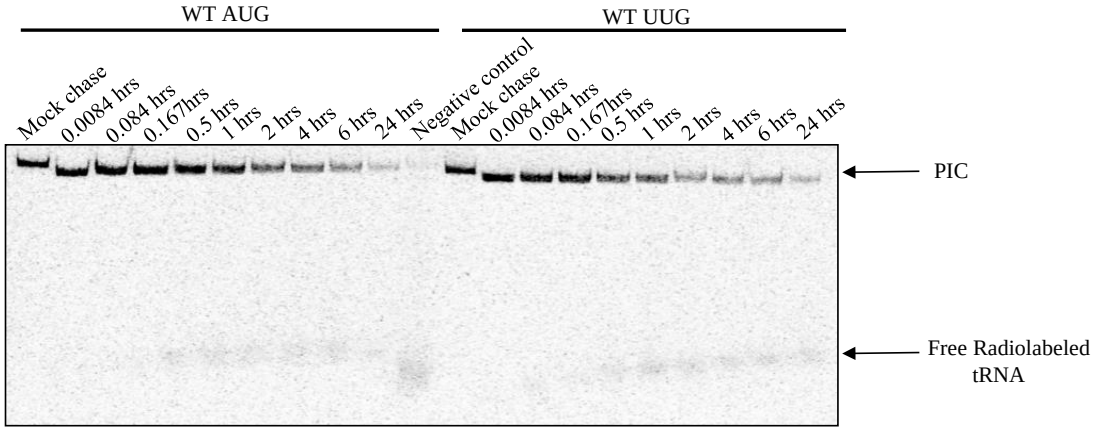
**(A-E)** To measure TC dissociation kinetics, as summarized schematically in Fig. 3.10.5 D, partial 48S complexes were assembled with eIF1A, model mRNA containing an AUG or UUG start codon, radiolabeled TC containing either WT or mutant eIF2β, and either WT or mutant eIF1. Following incubation at 26° C for two hours, each reaction was chased with excess unlabeled TC for increasing periods of time and the fraction of labeled Met-tRNA$_i$ bound to the PIC at each time-point was determined via EMSA. Data from each gel were plotted individually to determine a dissociation rate ($k_{off}$). Representative plots are shown in Figs. 3.10.5 E, 3.10.5 H, 3.10.7 C, 3.10.8 G, and 3.10.9 E. The rates from three independent experiments were averaged to determine the $k_{off}$ values shown in Figs. 3.10.5 F, 3.10.5 I, 3.10.7 D, 3.10.8 H, and 3.10.9 F. In order to give a visual representation of the variability across gels, we have averaged together values at each time point across three independent experiments and plotted an average curve with SEMs for eIF2β-F217A/Q221A (A), eIF1-F108D (B), eIF1-Q31E (C), eIF1-F108R (D), eIF2β-S202A/K214A (E), and eIF2β-E189R (F).

Fig. S5

A


AUG
TC
+
40S
1A
1
1A
1
A U G
1A
1
A U G
1A
1
A U G

B


WT AUG
WT UUG
Mock chase
0.0084 hrs
0.084 hrs
0.167hrs
0.5 hrs
1 hrs
2 hrs
4 hrs
6 hrs
24 hrs
Negative control
Mock chase
0.0084 hrs
0.084 hrs
0.167hrs
0.5 hrs
1 hrs
2 hrs
4 hrs
6 hrs
24 hrs
PIC
Free Radiolabeled tRNA

C

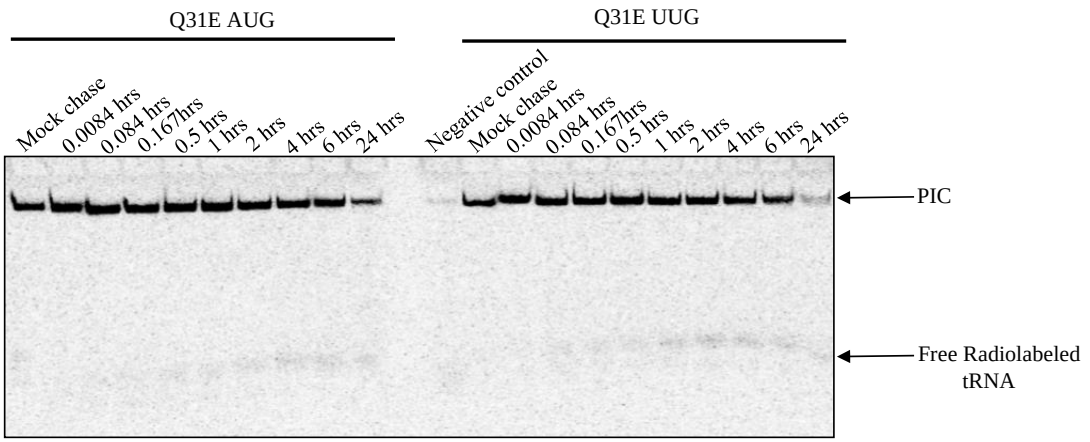
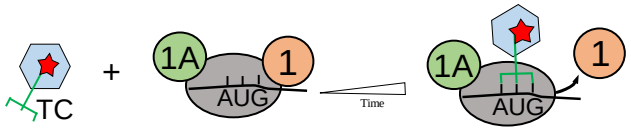

Q31E AUG
Q31E UUG
Mock chase
0.0084 hrs
0.084 hrs
0.167hrs
0.5 hrs
1 hrs
2 hrs
4 hrs
6 hrs
24 hrs
Negative control
Mock chase
0.0084 hrs
0.084 hrs
0.167hrs
0.5 hrs
1 hrs
2 hrs
4 hrs
6 hrs
24 hrs
PIC
Free Radiolabeled tRNA

**Figure 7.2.5. Representative results from TC dissociation assays.**

**(A-C)** To measure TC dissociation kinetics, as summarized schematically in (A), partial 48S complexes were assembled with radiolabeled TC, eIF1A, model mRNA containing an AUG or UUG start codon, and either WT eIF1 (B) or eIF1-Q31E (C). Following incubation at 26° C for two hours, each reaction was chased with excess unlabeled TC for increasing periods of time (between 0.0084 and 24 hours as indicated above gel image) and the fraction of labeled Met-$tRNA_i$ bound to the PIC at each time-point was determined by resolving radiolabeled 48S complexes (upper band) from free radiolabeled $tRNA_i$ (lower band) by EMSA. As a control, one reaction was chased with buffer only ('mock chase') for each eIF1/mRNA pair, representing the maximum possible PIC-bound radioactivity. One reaction was also conducted in which unlabeled chase was added before labeled TC ('negative control'), demonstrating the least possible PIC-bound radioactivity. The upper band from each lane was quantified and normalized to total radioactivity in the lane (after background subtraction). These values for each time point were normalized to the negative control and plotted as shown in Fig. 3.10.5 H.

Fig. S6

A


TC
+
1A
1
AUG
Time
1A
1
AUG

B

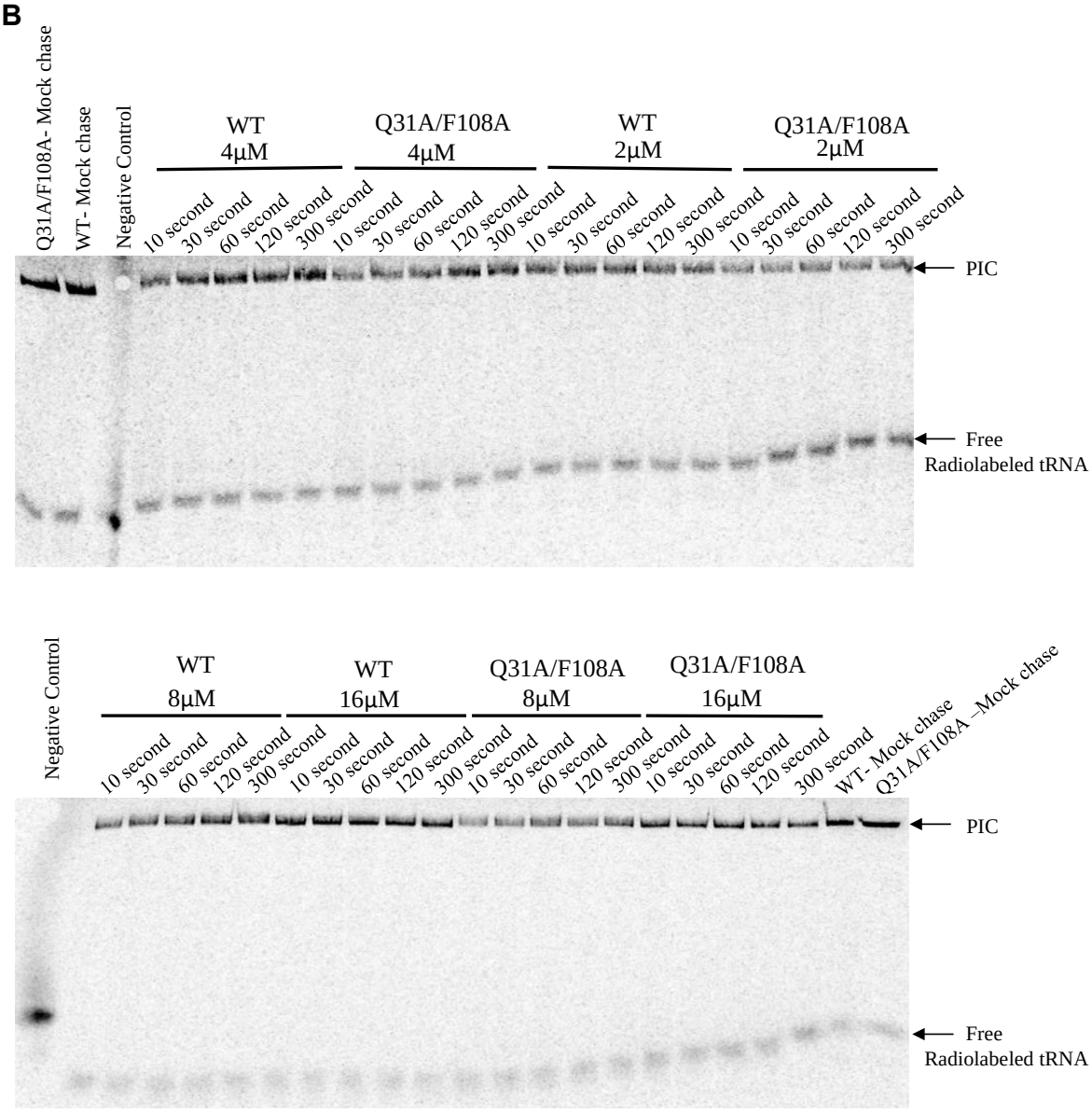

Q31A/F108A- Mock chase
WT- Mock chase
Negative Control
WT 4μM
Q31A/F108A 4μM
WT 2μM
Q31A/F108A 2μM
10 second
30 second
60 second
120 second
300 second
PIC
Free Radiolabeled tRNA
Negative Control
WT 8μM
WT 16μM
Q31A/F108A 8μM
Q31A/F108A 16μM
WT- Mock chase
Q31A/F108A –Mock chase
PIC
Free Radiolabeled tRNA

**Figure 7.2.6. Representative results from TC association assays.**

**(A, B)** To measure TC association kinetics, as summarized schematically in (A), radiolabeled TC was mixed with pre-formed 40S·eIF1A·eIF1·mRNA complexes containing either WT eIF1 or eIF1-Q31A/F108A and incubated for increasing times (as indicated). Reactions for WT and mutant eIF1 were carried out at each of four concentrations of 40S ribosomal subunits (2 μM, 4 μM, 8 μM, and 16 μM, as shown). Reactions were terminated with a chase of excess unlabeled TC. The fraction of labeled Met-tRNA$_i$ bound to the PIC at each time-point was determined by EMSA, as in Fig. 7.2.5. As a control, one reaction was chased with buffer only ('mock chase'), representing the maximum possible PIC-bound radioactivity. One reaction was also carried out in which unlabeled chase was added before labeled TC ('negative control'), demonstrating the least possible PIC-bound radioactivity. For each time point, the upper band was quantified and normalized to total radioactivity in the lane (after background subtraction), followed by normalization to the negative control. These values were plotted to obtain the pseudo-first-order rate constant ($k_{obs}$) at each 40S concentration. The resulting $k_{obs}$ values were plotted versus 40S concentration, as in Fig. 3.10.6 B, to obtain the second-order rate constant ($k_{on}$).

Fig. S7

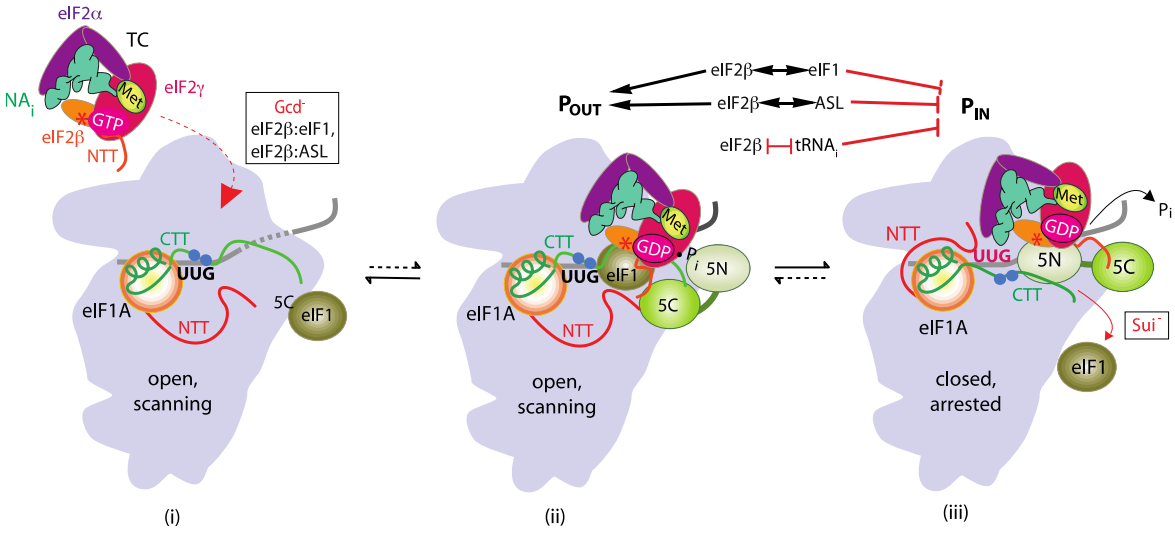

eIF2α
TC
eIF2γ
Met
GTP
eIF2β
NTT
NA$_i$
Gcd$^-$
eIF2β:eIF1,
eIF2β:ASL
CTT
UUG
eIF1A
5C
eIF1
NTT
open,
scanning
(i)
eIF2β
eIF1
P$_{OUT}$
eIF2β
ASL
P$_{IN}$
eIF2β
tRNA$_i$
Met
GDp
P$_i$
5N
CTT
UUG
eIF1
5C
eIF1A
NTT
open,
scanning
(ii)
Met
GDp
P$_i$
NTT
UUG
5N
5C
CTT
eIF1A
Sui$^-$
eIF1
closed,
arrested
(iii)

**Figure 7.2.7. Model describing the effects of eIF2β interactions on the conformational rearrangements of the PIC during start codon selection.**

As described in Fig. 3.10.1 A, eIF1 and eIF1A promote an open, scanning conformation of the PIC to which TC loads (i). Mutations that destabilize the open conformation slow TC loading, resulting in a Gcd$^-$ phenotype. Following TC binding, the PIC scans the mRNA leader in the open conformation (ii). When a start codon is recognized, the PIC transitions to the closed state, accompanied by release of eIF1 and $P_i$ (iii). Mutations that favor the closed complex, either by destabilizing the open state or by removing an impediment to transition to the closed state, decrease the fidelity of initiation, allowing selection of non-AUG codons including the near cognate UUG. Such mutations confer a Sui$^-$ phenotype. (Above) The arrows summarize the contributions of eIF2β to ensuring accurate start codon selection indicated by our findings here. The interactions of eIF2β with eIF1 and the ASL of $tRNA_i$ (Fig. 7.2.1 D) stabilize and promote the open state, either by anchoring eIF1 to the complex or by stabilizing $tRNA_i$ binding, while impeding transition to the closed state. For this reason, substitutions at these interfaces confer both Gcd$^-$ and Sui$^-$ phenotypes. The clash of open-state eIF2β with the closed- state $tRNA_i$ D-loop (Fig. 7.2.1 C), however, enforces a strict requirement for an AUG start codon and inhibits transition to the closed state in its absence, but does not stabilize the open state. Thus, substitutions at this interface produce Sui$^-$ phenotypes without any accompanying Gcd$^-$ phenotype.

# 8. REFERENCES


1. Kohler R, Mooney RA, Mills DJ, Landick R, & Cramer P (2017) Architecture of a transcribing-translating expressome. *Science* 356(6334):194-197.
2. Lee CP, Seong BL, & RajBhandary UL (1991) Structural and sequence elements important for recognition of Escherichia coli formylmethionine tRNA by methionyl-tRNA transformylase are clustered in the acceptor stem. *J Biol Chem* 266(27):18012-18017.
3. Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* 361:13-37.
4. Dahlquist KD & Puglisi JD (2000) Interaction of translation initiation factor IF1 with the E. coli ribosomal A site. *J Mol Biol* 299(1):1-15.
5. Boelens R & Gualerzi CO (2002) Structure and function of bacterial initiation factors. *Curr Protein Pept Sci* 3(1):107-119.
6. Petersen HU, Roll T, Grunberg-Manago M, & Clark BF (1979) Specific interaction of initiation factor IF2 of E. coli with formylmethionyl-tRNA f Met. *Biochem Biophys Res Commun* 91(3):1068-1074.
7. Maracci C & Rodnina MV (2016) Review: Translational GTPases. *Biopolymers* 105(8):463-475.
8. Milon P*, et al.* (2010) The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *EMBO Rep* 11(4):312-316.
9. Grunberg-Manago M*, et al.* (1975) Light-scattering studies showing the effect of initiation factors on the reversible dissociation of Escherichia coli ribosomes. *J Mol Biol* 94(3):461-478.
10. Lomakin IB, Shirokikh NE, Yusupov MM, Hellen CU, & Pestova TV (2006) The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J* 25(1):196-210.
11. Rodnina MV (2018) Translation in Prokaryotes. *Cold Spring Harb Perspect Biol* 10(9):a032664.
12. Hinnebusch AG (2014) The scanning mechanism of eukaryotic translation initiation. *Annual review of biochemistry* 83:779-812.
13. James CC & Smyth JW (2018) Alternative mechanisms of translation initiation: An emerging dynamic regulator of the proteome in health and disease. *Life Sci* 212:138-144.
14. Roberts L & Wieden HJ (2018) Viruses, IRESs, and a universal translation initiation mechanism. *Biotechnol Genet Eng Rev* 34(1):60-75.
15. Pelletier J & Sonenberg N (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334(6180):320-325.
16. Sarnow P (1989) Translation of glucose-regulated protein 78/immunoglobulin heavy-chain binding protein mRNA is increased in poliovirus-infected cells at a time when cap-dependent translation of cellular mRNAs is inhibited. *Proc Natl Acad Sci U S A* 86(15):5795-5799.
17. Shi Y, Sharma A, Wu H, Lichtenstein A, & Gera J (2005) Cyclin D1 and c-myc internal ribosome entry site (IRES)-dependent translation is regulated by AKT

activity and enhanced by rapamycin through a p38 MAPK- and ERK-dependent pathway. *J Biol Chem* 280(12):10964-10973.
18. Stein I, *et al.* (1998) Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia. *Mol Cell Biol* 18(6):3112-3119.
19. Stoneley M, Paulin FE, Le Quesne JP, Chappell SA, & Willis AE (1998) C-Myc 5' untranslated region contains an internal ribosome entry segment. *Oncogene* 16(3):423-428.
20. Weingarten-Gabbay S, *et al.* (2016) Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* 351(6270).
21. Mokrejs M, *et al.* (2006) IRESite: the database of experimentally verified IRES structures (www.iresite.org). *Nucleic Acids Res* 34(Suppl_1):D125-130.
22. Pamudurti NR, *et al.* (2017) Translation of CircRNAs. *Mol Cell* 66(1):9-21 e27.
23. Kwan T & Thompson SR (2019) Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb Perspect Biol* 11(4):a032672.
24. Green KM, Linsalata AE, & Todd PK (2016) RAN translation-What makes it run? *Brain Res* 1647:30-42.
25. Kearse MG & Wilusz JE (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev* 31(17):1717-1731.
26. Lasko P (2009) Translational control during early development. *Prog Mol Biol Transl Sci* 90:211-254.
27. Richter JD & Lasko P (2011) Translational control in oocyte development. *Cold Spring Harb Perspect Biol* 3(9):a002758.
28. Sonenberg N & Hinnebusch AG (2007) New modes of translational control in development, behavior, and disease. *Mol Cell* 28(5):721-729.
29. Nakahata Y & Yasuda R (2018) Plasticity of Spine Structure: Local Signaling, Translation and Cytoskeletal Reorganization. *Front Synaptic Neurosci* 10:29.
30. Miller S, *et al.* (2002) Disruption of dendritic translation of CaMKIIalpha impairs stabilization of synaptic plasticity and memory consolidation. *Neuron* 36(3):507-519.
31. Costa-Mattioli M, *et al.* (2007) eIF2alpha phosphorylation bidirectionally regulates the switch from short- to long-term synaptic plasticity and memory. *Cell* 129(1):195-206.
32. Delaidelli A, Jan A, Herms J, & Sorensen PH (2019) Translational control in brain pathologies: biological significance and therapeutic opportunities. *Acta Neuropathol* 137(4):535-555.
33. Hagerman RJ, *et al.* (2017) Fragile X syndrome. *Nat Rev Dis Primers* 3:17065.
34. Napoli I, *et al.* (2008) The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* 134(6):1042-1054.
35. Bugiani M, Vuong C, Breur M, & van der Knaap MS (2018) Vanishing white matter: a leukodystrophy due to astrocytic dysfunction. *Brain Pathol* 28(3):408-421.
36. Moon SL, Sonenberg N, & Parker R (2018) Neuronal Regulation of eIF2alpha Function in Health and Neurological Disorders. *Trends Mol Med* 24(6):575-589.

37. Young-Baird SK, Shin BS, & Dever TE (2019) MEHMO syndrome mutation EIF2S3-I259M impairs initiator Met-tRNAiMet binding to eukaryotic translation initiation factor eIF2. *Nucleic Acids Res* 47(2):855-867.
38. Hernandez-Ortega K, Garcia-Esparcia P, Gil L, Lucas JJ, & Ferrer I (2016) Altered Machinery of Protein Synthesis in Alzheimer's: From the Nucleolus to the Ribosome. *Brain Pathol* 26(5):593-605.
39. Taymans JM, Nkiliza A, & Chartier-Harlin MC (2015) Deregulation of protein translation control, a potential game-changing hypothesis for Parkinson's disease pathogenesis. *Trends Mol Med* 21(8):466-472.
40. Deng H, Wu Y, & Jankovic J (2015) The EIF4G1 gene and Parkinson's disease. *Acta Neurol Scand* 132(2):73-78.
41. An WL*, et al.* (2003) Up-regulation of phosphorylated/activated p70 S6 kinase and its relationship to neurofibrillary pathology in Alzheimer's disease. *Am J Pathol* 163(2):591-607.
42. Li X, Alafuzoff I, Soininen H, Winblad B, & Pei JJ (2005) Levels of mTOR and its downstream targets 4E-BP1, eEF2, and eEF2 kinase in relationships with tau in Alzheimer's disease brain. *FEBS J* 272(16):4211-4220.
43. Garcia-Esparcia P*, et al.* (2015) Altered machinery of protein synthesis is region- and stage-dependent and is associated with alpha-synuclein oligomers in Parkinson's disease. *Acta Neuropathol Commun* 3:76.
44. Harding HP & Ron D (2002) Endoplasmic reticulum stress and the development of diabetes: a review. *Diabetes* 51 Suppl 3:S455-461.
45. Delepine M*, et al.* (2000) EIF2AK3, encoding translation initiation factor 2-alpha kinase 3, is mutated in patients with Wolcott-Rallison syndrome. *Nat Genet* 25(4):406-409.
46. Pause A*, et al.* (1994) Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 371(6500):762-767.
47. Proud CG*, et al.* (2001) Interplay between insulin and nutrients in the regulation of translation factors. *Biochem Soc Trans* 29(Pt 4):541-547.
48. Kostyak JC, Kimball SR, Jefferson LS, & Farrell PA (2001) Severe diabetes inhibits resistance exercise-induced increase in eukaryotic initiation factor 2B activity. *J Appl Physiol (1985)* 91(1):79-84.
49. Pende M*, et al.* (2000) Hypoinsulinaemia, glucose intolerance and diminished beta-cell size in S6K1-deficient mice. *Nature* 408(6815):994-997.
50. Shi Y, Taylor SI, Tan SL, & Sonenberg N (2003) When translation meets metabolism: multiple links to diabetes. *Endocr Rev* 24(1):91-101.
51. Ruggero D (2013) Translational control in cancer etiology. *Cold Spring Harb Perspect Biol* 5(2):a012336.
52. Silvera D, Formenti SC, & Schneider RJ (2010) Translational control in cancer. *Nat Rev Cancer* 10(4):254-266.
53. Hannan KM, Sanij E, Hein N, Hannan RD, & Pearson RB (2011) Signaling to the ribosome in cancer--It is more than just mTORC1. *IUBMB Life* 63(2):79-85.
54. Sorrells DL*, et al.* (1998) Detection of eIF4E gene amplification in breast cancer by competitive PCR. *Ann Surg Oncol* 5(3):232-237.

55. Lazaris-Karatzas A, Montine KS, & Sonenberg N (1990) Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature* 345(6275):544-547.
56. De Benedetti A & Rhoads RE (1990) Overexpression of eukaryotic protein synthesis initiation factor 4E in HeLa cells results in aberrant growth and morphology. *Proc Natl Acad Sci U S A* 87(21):8212-8216.
57. Graff JR*, et al.* (2007) Therapeutic suppression of translation initiation factor eIF4E expression reduces tumor growth without toxicity. *J Clin Invest* 117(9):2638-2648.
58. Wada H, Ivester CT, Carabello BA, Cooper Gt, & McDermott PJ (1996) Translational initiation factor eIF-4E. A link between cardiac load and protein synthesis. *J Biol Chem* 271(14):8359-8364.
59. Tuxworth WJ, Jr., Saghir AN, Spruill LS, Menick DR, & McDermott PJ (2004) Regulation of protein synthesis by eIF4E phosphorylation in adult cardiocytes: the consequence of secondary structure in the 5'-untranslated region of mRNA. *Biochem J* 378(Pt 1):73-82.
60. Bandhuvula P*, et al.* (2011) S1P lyase: a novel therapeutic target for ischemia-reperfusion injury of the heart. *Am J Physiol Heart Circ Physiol* 300(5):H1753-1761.
61. Zhang D*, et al.* (2010) MTORC1 regulates cardiac function and myocyte survival through 4E-BP1 inhibition in mice. *J Clin Invest* 120(8):2805-2816.
62. Scheper GC, van der Knaap MS, & Proud CG (2007) Translation matters: protein synthesis defects in inherited disease. *Nat Rev Genet* 8(9):711-723.
63. Mills EW & Green R (2017) Ribosomopathies: There's strength in numbers. *Science* 358(6363).
64. Tahmasebi S, Khoutorsky A, Mathews MB, & Sonenberg N (2018) Translation deregulation in human disease. *Nat Rev Mol Cell Biol* 19(12):791-807.
65. Simsek D*, et al.* (2017) The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* 169(6):1051-1065 e1018.
66. Allerson CR, Cazzola M, & Rouault TA (1999) Clinical severity and thermodynamic effects of iron-responsive element mutations in hereditary hyperferritinemia-cataract syndrome. *J Biol Chem* 274(37):26439-26447.
67. Kondo T*, et al.* (1998) Familial essential thrombocythemia associated with one-base deletion in the 5'-untranslated region of the thrombopoietin gene. *Blood* 92(4):1091-1096.
68. Liu L*, et al.* (1999) Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma. *Nat Genet* 21(1):128-132.
69. Claude A (1944) The Constitution of Mitochondria and Microsomes, and the Distribution of Nucleic Acid in the Cytoplasm of a Leukemic Cell. *J Exp Med* 80(1):19-29.
70. Palade GE (1955) A small particulate component of the cytoplasm. *J Biophys Biochem Cytol* 1(1):59-68.
71. Littlefield JW, Keller EB, Gross J, & Zamecnik PC (1955) Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat. *J Biol Chem* 217(1):111-123.
72. Voet D & Voet JG (2004) *Biochemistry* (J. Wiley & Sons, New York) 3rd Ed.

73. Jobe A, Liu Z, Gutierrez-Vargas C, & Frank J (2019) New Insights into Ribosome Structure and Function. *Cold Spring Harb Perspect Biol* 11(1):a032615.
74. Wilson DN & Doudna Cate JH (2012) The structure and function of the eukaryotic ribosome. *Cold Spring Harb Perspect Biol* 4(5):a011536.
75. Wimberly BT*, et al.* (2000) Structure of the 30S ribosomal subunit. *Nature* 407(6802):327-339.
76. Schluenzen F*, et al.* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* 102(5):615-623.
77. Ban N, Nissen P, Hansen J, Moore PB, & Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 289(5481):905-920.
78. Harms J*, et al.* (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* 107(5):679-688.
79. Yusupov MM*, et al.* (2001) Crystal structure of the ribosome at 5.5 A resolution. *Science* 292(5518):883-896.
80. Schuwirth BS*, et al.* (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science* 310(5749):827-834.
81. Selmer M*, et al.* (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313(5795):1935-1942.
82. Spahn CM*, et al.* (2001) Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions. *Cell* 107(3):373-386.
83. Armache JP*, et al.* (2010) Localization of eukaryote-specific ribosomal proteins in a 5.5-A cryo-EM map of the 80S eukaryotic ribosome. *Proc Natl Acad Sci U S A* 107(46):19754-19759.
84. Sengupta J*, et al.* (2004) Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nat Struct Mol Biol* 11(10):957-962.
85. Halic M, Becker T, Frank J, Spahn CM, & Beckmann R (2005) Localization and dynamic behavior of ribosomal protein L30e. *Nat Struct Mol Biol* 12(5):467-468.
86. Chandramouli P*, et al.* (2008) Structure of the mammalian 80S ribosome at 8.7 A resolution. *Structure* 16(4):535-548.
87. Taylor DJ*, et al.* (2009) Comprehensive molecular structure of the eukaryotic ribosome. *Structure* 17(12):1591-1604.
88. Schuler M*, et al.* (2006) Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nat Struct Mol Biol* 13(12):1092-1096.
89. Becker T*, et al.* (2009) Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science* 326(5958):1369-1373.
90. Klinge S, Voigts-Hoffmann F, Leibundgut M, Arpagaus S, & Ban N (2011) Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* 334(6058):941-948.
91. Rabl J, Leibundgut M, Ataide SF, Haag A, & Ban N (2011) Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 331(6018):730-736.
92. Ben-Shem A*, et al.* (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. *Science* 334(6062):1524-1529.

93. Voorhees RM & Hegde RS (2016) Structure of the Sec61 channel opened by a signal sequence. *Science* 351(6268):88-91.
94. Voorhees RM, Fernandez IS, Scheres SH, & Hegde RS (2014) Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. *Cell* 157(7):1632-1643.
95. Khatter H, Myasnikov AG, Natchiar SK, & Klaholz BP (2015) Structure of the human 80S ribosome. *Nature* 520(7549):640-645.
96. Hashem Y*, et al.* (2013) Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* 503(7477):539-543.
97. Fernandez IS, Bai XC, Murshudov G, Scheres SH, & Ramakrishnan V (2014) Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* 157(4):823-831.
98. Yamamoto H*, et al.* (2015) Molecular architecture of the ribosome-bound Hepatitis C Virus internal ribosomal entry site RNA. *EMBO J* 34(24):3042-3058.
99. Yamamoto H*, et al.* (2014) Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat Struct Mol Biol* 21(8):721-727.
100. Muhs M*, et al.* (2015) Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. *Mol Cell* 57(3):422-432.
101. Quade N, Boehringer D, Leibundgut M, van den Heuvel J, & Ban N (2015) Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-A resolution. *Nat Commun* 6:7646.
102. Abeyrathne PD, Koh CS, Grant T, Grigorieff N, & Korostelev AA (2016) Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome. *Elife* 5:e14874.
103. Murray J*, et al.* (2016) Structural characterization of ribosome recruitment and translocation by type IV IRES. *Elife* 5:e13567.
104. Hashem Y*, et al.* (2013) High-resolution cryo-electron microscopy structure of the Trypanosoma brucei ribosome. *Nature* 494(7437):385-389.
105. Shalev-Benami M*, et al.* (2016) 2.8-A Cryo-EM Structure of the Large Ribosomal Subunit from the Eukaryotic Parasite Leishmania. *Cell Rep* 16(2):288-294.
106. Zhang X*, et al.* (2016) Structures and stabilization of kinetoplastid-specific split rRNAs revealed by comparing leishmanial and human ribosomes. *Nat Commun* 7:13223.
107. Liu Z*, et al.* (2016) Structure and assembly model for the Trypanosoma cruzi 60S ribosomal subunit. *Proc Natl Acad Sci U S A* 113(43):12174-12179.
108. Liu Z*, et al.* (2017) Determination of the ribosome structure to a resolution of 2.5 A by single-particle cryo-EM. *Protein Sci* 26(1):82-92.
109. Hashem Y*, et al.* (2013) Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* 153(5):1108-1119.
110. Hussain T*, et al.* (2014) Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* 159(3):597-607.
111. Llacer JL*, et al.* (2015) Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Mol Cell* 59(3):399-412.
112. Llacer JL*, et al.* (2018) Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *Elife* 7:e39273.

113. Simonetti A, *et al.* (2016) eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. *Mol Cell* 63(2):206-217.
114. Aylett CH, Boehringer D, Erzberger JP, Schaefer T, & Ban N (2015) Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nat Struct Mol Biol* 22(3):269-271.
115. Fernandez IS, *et al.* (2013) Molecular architecture of a eukaryotic translational initiation complex. *Science* 342(6160):1240585.
116. des Georges A, *et al.* (2015) Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* 525(7570):491-495.
117. Hinnebusch AG (2017) Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci* 42(8):589-611.
118. Ban N, *et al.* (2014) A new system for naming ribosomal proteins. *Curr Opin Struct Biol* 24:165-169.
119. Dong J, *et al.* (2017) Rps3/uS3 promotes mRNA binding at the 40S ribosome entry channel and stabilizes preinitiation complexes at start codons. *Proc Natl Acad Sci U S A* 114(11):E2126-E2135.
120. Gunisova S, Hronova V, Mohammad MP, Hinnebusch AG, & Valasek LS (2018) Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiol Rev* 42(2):165-192.
121. Paulin FE, Campbell LE, O'Brien K, Loughlin J, & Proud CG (2001) Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Curr Biol* 11(1):55-59.
122. Merrick WC (1992) Mechanism and regulation of eukaryotic protein synthesis. *Microbiol Rev* 56(2):291-315.
123. Kapp LD & Lorsch JR (2004) GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *J Mol Biol* 335(4):923-936.
124. Astrom SU, von Pawel-Rammingen U, & Bystrom AS (1993) The yeast initiator tRNAMet can act as an elongator tRNA(Met) in vivo. *J Mol Biol* 233(1):43-58.
125. Farruggio D, Chaudhuri J, Maitra U, & RajBhandary UL (1996) The A1 x U72 base pair conserved in eukaryotic initiator tRNAs is important specifically for binding to the eukaryotic translation initiation factor eIF2. *Mol Cell Biol* 16(8):4248-4256.
126. Kapp LD, Kolitz SE, & Lorsch JR (2006) Yeast initiator tRNA identity elements cooperate to influence multiple steps of translation initiation. *RNA* 12(5):751-764.
127. Passmore LA, *et al.* (2007) The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* 26(1):41-50.
128. Olsen DS, *et al.* (2003) Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. *EMBO J* 22(2):193-204.
129. Cheung YN, *et al.* (2007) Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. *Genes Dev* 21(10):1217-1230.
130. Algire MA, *et al.* (2002) Development and characterization of a reconstituted yeast translation initiation system. *RNA* 8(3):382-397.
131. Maag D & Lorsch JR (2003) Communication between eukaryotic translation initiation factors 1 and 1A on the yeast small ribosomal subunit. *J Mol Biol* 330(5):917-924.

132. Sokabe M & Fraser CS (2014) Human eukaryotic initiation factor 2 (eIF2)-GTP-Met-tRNAi ternary complex and eIF3 stabilize the 43 S preinitiation complex. *J Biol Chem* 289(46):31827-31836.
133. Majumdar R, Bandyopadhyay A, & Maitra U (2003) Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *J Biol Chem* 278(8):6580-6587.
134. Kolupaeva VG, Unbehaun A, Lomakin IB, Hellen CU, & Pestova TV (2005) Binding of eukaryotic initiation factor 3 to ribosomal 40S subunits and its role in ribosomal dissociation and anti-association. *RNA* 11(4):470-486.
135. Valasek LS*, et al.* (2017) Embraced by eIF3: structural and functional insights into the roles of eIF3 across the translation cycle. *Nucleic Acids Res* 45(19):10948-10968.
136. Aitken CE*, et al.* (2016) Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. *Elife* 5:e20934.
137. Maag D, Fekete CA, Gryczynski Z, & Lorsch JR (2005) A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Mol Cell* 17(2):265-275.
138. Weisser M, Voigts-Hoffmann F, Rabl J, Leibundgut M, & Ban N (2013) The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat Struct Mol Biol* 20(8):1015-1017.
139. Asano K, Clayton J, Shalev A, & Hinnebusch AG (2000) A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. *Genes Dev* 14(19):2534-2546.
140. Dennis MD, Person MD, & Browning KS (2009) Phosphorylation of plant translation initiation factors by CK2 enhances the in vitro interaction of multifactor complex components. *J Biol Chem* 284(31):20615-20628.
141. Sokabe M, Fraser CS, & Hershey JW (2012) The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. *Nucleic Acids Res* 40(2):905-913.
142. Merrick WC & Pavitt GD (2018) Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb Perspect Biol* 10(12):a033092.
143. Luna RE*, et al.* (2012) The C-terminal domain of eukaryotic initiation factor 5 promotes start codon recognition by its dynamic interplay with eIF1 and eIF2beta. *Cell Rep* 1(6):689-702.
144. Yamamoto Y*, et al.* (2005) The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G. *Proc Natl Acad Sci U S A* 102(45):16164-16169.
145. Valasek L, Nielsen KH, & Hinnebusch AG (2002) Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. *EMBO J* 21(21):5886-5898.
146. Valasek L, Nielsen KH, Zhang F, Fekete CA, & Hinnebusch AG (2004) Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. *Mol Cell Biol* 24(21):9437-9455.

147. Singh CR, Yamamoto Y, & Asano K (2004) Physical association of eukaryotic initiation factor (eIF) 5 carboxyl-terminal domain with the lysine-rich eIF2beta segment strongly enhances its binding to eIF3. *J Biol Chem* 279(48):49644-49655.
148. Gross JD*, et al.* (2003) Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115(6):739-750.
149. Blum S*, et al.* (1992) ATP hydrolysis by initiation factor 4A is required for translation initiation in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* 89(16):7664-7668.
150. Svitkin YV*, et al.* (2001) The requirement for eukaryotic initiation factor 4A (elF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA* 7(3):382-394.
151. Pestova TV & Kolupaeva VG (2002) The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* 16(22):2906-2922.
152. Mitchell SF*, et al.* (2010) The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. *Mol Cell* 39(6):950-962.
153. Sen ND, Zhou F, Ingolia NT, & Hinnebusch AG (2015) Genome-wide analysis of translational efficiency reveals distinct but overlapping functions of yeast DEAD-box RNA helicases Ded1 and eIF4A. *Genome Res* 25(8):1196-1205.
154. Yourik P*, et al.* (2017) Yeast eIF4A enhances recruitment of mRNAs regardless of their structural complexity. *Elife* 6:e31476.
155. Sokabe M & Fraser CS (2017) A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Proc Natl Acad Sci U S A* 114(24):6304-6309.
156. Dmitriev SE, Terenin IM, Dunaevsky YE, Merrick WC, & Shatsky IN (2003) Assembly of 48S translation initiation complexes from purified components with mRNAs that have some base pairing within their 5' untranslated regions. *Mol Cell Biol* 23(24):8925-8933.
157. Shahbazian D*, et al.* (2010) Control of cell survival and proliferation by mammalian eukaryotic initiation factor 4B. *Mol Cell Biol* 30(6):1478-1485.
158. Walker SE*, et al.* (2013) Yeast eIF4B binds to the head of the 40S ribosomal subunit and promotes mRNA recruitment through its N-terminal and internal repeat domains. *RNA* 19(2):191-207.
159. Sen ND, Zhou F, Harris MS, Ingolia NT, & Hinnebusch AG (2016) eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc Natl Acad Sci U S A* 113(38):10464-10472.
160. Korneeva NL, Lamphear BJ, Hennigan FL, & Rhoads RE (2000) Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J Biol Chem* 275(52):41369-41376.
161. LeFebvre AK*, et al.* (2006) Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem* 281(32):22917-22932.

162. Yanagiya A*, et al.* (2009) Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol* 29(6):1661-1669.
163. Villa N, Do A, Hershey JW, & Fraser CS (2013) Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J Biol Chem* 288(46):32932-32940.
164. Park EH*, et al.* (2011) Multiple elements in the eIF4G1 N-terminus promote assembly of eIF4G1*PABP mRNPs in vivo. *EMBO J* 30(2):302-316.
165. Wells SE, Hillner PE, Vale RD, & Sachs AB (1998) Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 2(1):135-140.
166. Archer SK, Shirokikh NE, Beilharz TH, & Preiss T (2016) Dynamics of ribosome scanning and recycling revealed by translation complex profiling. *Nature* 535(7613):570-574.
167. Archer SK, Shirokikh NE, Hallwirth CV, Beilharz TH, & Preiss T (2015) Probing the closed-loop model of mRNA translation in living cells. *RNA Biol* 12(3):248-254.
168. Costello J*, et al.* (2015) Global mRNA selection mechanisms for translation initiation. *Genome Biol* 16:10.
169. Costello JL*, et al.* (2017) Dynamic changes in eIF4F-mRNA interactions revealed by global analyses of environmental stress responses. *Genome Biol* 18(1):201.
170. Kumar P, Hellen CU, & Pestova TV (2016) Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev* 30(13):1573-1588.
171. Phan L, Schoenfeld LW, Valasek L, Nielsen KH, & Hinnebusch AG (2001) A subcomplex of three eIF3 subunits binds eIF1 and eIF5 and stimulates ribosome binding of mRNA and tRNA(i)Met. *EMBO J* 20(11):2954-2965.
172. Jivotovskaya AV, Valasek L, Hinnebusch AG, & Nielsen KH (2006) Eukaryotic translation initiation factor 3 (eIF3) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. *Mol Cell Biol* 26(4):1355-1372.
173. Trachsel H & Staehelin T (1979) Initiation of mammalian protein synthesis. The multiple functions of the initiation factor eIF-3. *Biochim Biophys Acta* 565(2):305-314.
174. Benne R & Hershey JW (1978) The mechanism of action of protein synthesis initiation factors from rabbit reticulocytes. *J Biol Chem* 253(9):3078-3087.
175. Fraser CS, Berry KE, Hershey JW, & Doudna JA (2007) eIF3j is located in the decoding center of the human 40S ribosomal subunit. *Mol Cell* 26(6):811-819.
176. Lee AS, Kranzusch PJ, Doudna JA, & Cate JH (2016) eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* 536(7614):96-99.
177. Rajagopal V, Park EH, Hinnebusch AG, & Lorsch JR (2012) Specific domains in yeast translation initiation factor eIF4G strongly bias RNA unwinding activity of the eIF4F complex toward duplexes with 5'-overhangs. *J Biol Chem* 287(24):20301-20312.
178. Kozak M (1980) Role of ATP in binding and migration of 40S ribosomal subunits. *Cell* 22(2 Pt 2):459-467.

179. Spirin AS (2009) How does a scanning ribosomal particle move along the 5'-untranslated region of eukaryotic mRNA? Brownian Ratchet model. *Biochemistry* 48(45):10688-10692.
180. Matsuda D & Dreher TW (2006) Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA* 12(7):1338-1349.
181. Garcia-Garcia C, Frieda KL, Feoktistova K, Fraser CS, & Block SM (2015) RNA BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* 348(6242):1486-1488.
182. Vassilenko KS, Alekhina OM, Dmitriev SE, Shatsky IN, & Spirin AS (2011) Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res* 39(13):5555-5567.
183. Berthelot K, Muldoon M, Rajkowitsch L, Hughes J, & McCarthy JE (2004) Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Mol Microbiol* 51(4):987-1001.
184. Abaeva IS, Marintchev A, Pisareva VP, Hellen CU, & Pestova TV (2011) Bypassing of stems versus linear base-by-base inspection of mammalian mRNAs during ribosomal scanning. *EMBO J* 30(1):115-129.
185. Firczuk H*, et al.* (2013) An in vivo control map for the eukaryotic mRNA translation machinery. *Mol Syst Biol* 9:635.
186. Rogers GW, Jr., Richter NJ, & Merrick WC (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J Biol Chem* 274(18):12236-12244.
187. Pisareva VP, Pisarev AV, Komar AA, Hellen CU, & Pestova TV (2008) Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. *Cell* 135(7):1237-1250.
188. Chuang RY, Weaver PL, Liu Z, & Chang TH (1997) Requirement of the DEAD-Box protein ded1p for messenger RNA translation. *Science* 275(5305):1468-1471.
189. Chiu WL*, et al.* (2010) The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol Cell Biol* 30(18):4415-4434.
190. Gupta N, Lorsch JR, & Hinnebusch AG (2018) Yeast Ded1 promotes 48S translation pre-initiation complex assembly in an mRNA-specific and eIF4F-dependent manner. *Elife* 7:e38892.
191. Parsyan A*, et al.* (2009) The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc Natl Acad Sci U S A* 106(52):22217-22222.
192. Cigan AM, Feng L, & Donahue TF (1988) tRNAi(met) functions in directing the scanning ribosome to the start site of translation. *Science* 242(4875):93-97.
193. Kolitz SE, Takacs JE, & Lorsch JR (2009) Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation. *RNA* 15(1):138-152.
194. Dever TE, Kinzy TG, & Pavitt GD (2016) Mechanism and Regulation of Protein Synthesis in Saccharomyces cerevisiae. *Genetics* 203(1):65-107.
195. Lomakin IB & Steitz TA (2013) The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* 500(7462):307-311.

196. Saini AK, Nanda JS, Lorsch JR, & Hinnebusch AG (2010) Regulatory elements in eIF1A control the fidelity of start codon selection by modulating tRNA(i)(Met) binding to the ribosome. *Genes Dev* 24(1):97-110.
197. Dong J*, et al.* (2008) Genetic identification of yeast 18S rRNA residues required for efficient recruitment of initiator tRNA(Met) and AUG selection. *Genes Dev* 22(16):2242-2255.
198. Nemoto N*, et al.* (2010) Yeast 18 S rRNA is directly involved in the ribosomal response to stringent AUG selection during translation initiation. *J Biol Chem* 285(42):32200-32212.
199. Zhang F, Saini AK, Shin BS, Nanda J, & Hinnebusch AG (2015) Conformational changes in the P site and mRNA entry channel evoked by AUG recognition in yeast translation preinitiation complexes. *Nucleic Acids Res* 43(4):2293-2312.
200. Thakur A & Hinnebusch AG (2018) eIF1 Loop 2 interactions with Met-tRNAi control the accuracy of start codon selection by the scanning preinitiation complex. *Proc Natl Acad Sci U S A* 115(18):E4159-E4168.
201. Varshney U, Lee CP, & RajBhandary UL (1993) From elongator tRNA to initiator tRNA. *Proc Natl Acad Sci U S A* 90(6):2305-2309.
202. Mandal N, Mangroo D, Dalluge JJ, McCloskey JA, & Rajbhandary UL (1996) Role of the three consecutive G:C base pairs conserved in the anticodon stem of initiator tRNAs in initiation of protein synthesis in Escherichia coli. *RNA* 2(5):473-482.
203. Dong J*, et al.* (2014) Conserved residues in yeast initiator tRNA calibrate initiation accuracy by regulating preinitiation complex stability at the start codon. *Genes Dev* 28(5):502-520.
204. Basavappa R & Sigler PB (1991) The 3 A crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination. *EMBO J* 10(10):3105-3111.
205. Fletcher CM, Pestova TV, Hellen CU, & Wagner G (1999) Structure and interactions of the translation initiation factor eIF1. *EMBO J* 18(9):2631-2637.
206. Martin-Marcos P*, et al.* (2017) eIF1A residues implicated in cancer stabilize translation preinitiation complexes and favor suboptimal initiation sites in yeast. *Elife* 6:e31250.
207. Algire MA, Maag D, & Lorsch JR (2005) Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* 20(2):251-262.
208. Fijalkowska D*, et al.* (2017) eIF1 modulates the recognition of suboptimal translation initiation sites and steers gene expression via uORFs. *Nucleic Acids Res* 45(13):7997-8013.
209. Yoon HJ & Donahue TF (1992) The suil suppressor locus in Saccharomyces cerevisiae encodes a translation factor that functions during tRNA(iMet) recognition of the start codon. *Mol Cell Biol* 12(1):248-260.
210. Nanda JS*, et al.* (2009) eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *J Mol Biol* 394(2):268-285.
211. Nanda JS, Saini AK, Munoz AM, Hinnebusch AG, & Lorsch JR (2013) Coordinated movements of eukaryotic translation initiation factors eIF1, eIF1A, and eIF5 trigger phosphate release from eIF2 in response to start codon

recognition by the ribosomal preinitiation complex. *J Biol Chem* 288(8):5316-5329.
212. Martin-Marcos P, Cheung YN, & Hinnebusch AG (2011) Functional elements in initiation factors 1, 1A, and 2beta discriminate against poor AUG context and non-AUG start codons. *Mol Cell Biol* 31(23):4814-4831.
213. Martin-Marcos P*, et al.* (2013) beta-Hairpin loop of eukaryotic initiation factor 1 (eIF1) mediates 40 S ribosome binding to regulate initiator tRNA(Met) recruitment and accuracy of AUG selection in vivo. *J Biol Chem* 288(38):27546-27562.
214. Martin-Marcos P*, et al.* (2014) Enhanced eIF1 binding to the 40S ribosome impedes conformational rearrangements of the preinitiation complex and elevates initiation accuracy. *RNA* 20(2):150-167.
215. Ivanov IP, Loughran G, Sachs MS, & Atkins JF (2010) Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci U S A* 107(42):18056-18060.
216. Yu J & Marintchev A (2018) Comparative sequence and structure analysis of eIF1A and eIF1AD. *BMC Struct Biol* 18(1):11.
217. Battiste JL, Pestova TV, Hellen CU, & Wagner G (2000) The eIF1A solution structure reveals a large RNA-binding surface important for scanning function. *Mol Cell* 5(1):109-119.
218. Sette M*, et al.* (1997) The structure of the translational initiation factor IF1 from E.coli contains an oligomer-binding motif. *EMBO J* 16(6):1436-1443.
219. Carter AP*, et al.* (2001) Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* 291(5503):498-501.
220. Yu Y*, et al.* (2009) Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing. *Nucleic Acids Res* 37(15):5167-5182.
221. Fekete CA*, et al.* (2005) The eIF1A C-terminal domain promotes initiation complex assembly, scanning and AUG selection in vivo. *EMBO J* 24(20):3588-3601.
222. Maag D, Algire MA, & Lorsch JR (2006) Communication between eukaryotic translation initiation factors 5 and 1A within the ribosomal pre-initiation complex plays a role in start site selection. *J Mol Biol* 356(3):724-737.
223. Fekete CA*, et al.* (2007) N- and C-terminal residues of eIF1A have opposing effects on the fidelity of start codon selection. *EMBO J* 26(6):1602-1614.
224. Luna RE*, et al.* (2013) The interaction between eukaryotic initiation factor 1A and eIF5 retains eIF1 within scanning preinitiation complexes. *Biochemistry* 52(52):9510-9518.
225. Conte MR*, et al.* (2006) Structure of the eukaryotic initiation factor (eIF) 5 reveals a fold common to several translation factors. *Biochemistry* 45(14):4550-4558.
226. Wei Z, Xue Y, Xu H, & Gong W (2006) Crystal structure of the C-terminal domain of S.cerevisiae eIF5. *J Mol Biol* 359(1):1-9.
227. Singh CR*, et al.* (2005) Eukaryotic translation initiation factor 5 is critical for integrity of the scanning preinitiation complex and accurate control of GCN4 translation. *Mol Cell Biol* 25(13):5480-5491.

228. Asano K*, et al.* (2001) A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNA(i)Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. *Cold Spring Harb Symp Quant Biol* 66:403-415.
229. Phan L*, et al.* (1998) Identification of a translation initiation factor 3 (eIF3) core complex, conserved in yeast and mammals, that interacts with eIF5. *Mol Cell Biol* 18(8):4935-4946.
230. Obayashi E*, et al.* (2017) Molecular Landscape of the Ribosome Pre-initiation Complex during mRNA Scanning: Structural Role for eIF3c and Its Control by eIF5. *Cell Rep* 18(11):2651-2663.
231. Loughran G, Sachs MS, Atkins JF, & Ivanov IP (2012) Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Res* 40(7):2898-2906.
232. Das S, Ghosh R, & Maitra U (2001) Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *J Biol Chem* 276(9):6720-6726.
233. Alone PV & Dever TE (2006) Direct binding of translation initiation factor eIF2gamma-G domain to its GTPase-activating and GDP-GTP exchange factors eIF5 and eIF2B epsilon. *J Biol Chem* 281(18):12636-12644.
234. Das S & Maitra U (2001) Functional significance and mechanism of eIF5-promoted GTP hydrolysis in eukaryotic translation initiation. *Prog Nucleic Acid Res Mol Biol* 70:207-231.
235. Huang HK, Yoon H, Hannig EM, & Donahue TF (1997) GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in Saccharomyces cerevisiae. *Genes Dev* 11(18):2396-2413.
236. Saini AK*, et al.* (2014) Eukaryotic translation initiation factor eIF5 promotes the accuracy of start codon recognition by regulating Pi release and conformational transitions of the preinitiation complex. *Nucleic Acids Res* 42(15):9623-9640.
237. Jennings MD & Pavitt GD (2010) eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. *Nature* 465(7296):378-381.
238. Jennings MD & Pavitt GD (2010) eIF5 is a dual function GAP and GDI for eukaryotic translational control. *Small GTPases* 1(2):118-123.
239. Jennings MD*, et al.* (2016) eIF2beta is critical for eIF5-mediated GDP-dissociation inhibitor activity and translational control. *Nucleic Acids Res* 44(20):9698-9709.
240. Singh CR*, et al.* (2006) An eIF5/eIF2 complex antagonizes guanine nucleotide exchange by eIF2B during translation initiation. *EMBO J* 25(19):4537-4546.
241. Singh CR*, et al.* (2007) Change in nutritional status modulates the abundance of critical pre-initiation intermediate complexes during translation initiation in vivo. *J Mol Biol* 370(2):315-330.
242. Jennings MD, Zhou Y, Mohammad-Qureshi SS, Bennett D, & Pavitt GD (2013) eIF2B promotes eIF5 dissociation from eIF2*GDP to facilitate guanine nucleotide exchange for translation initiation. *Genes Dev* 27(24):2696-2707.
243. Bogorad AM, Lin KY, & Marintchev A (2018) eIF2B Mechanisms of Action and Regulation: A Thermodynamic View. *Biochemistry* 57(9):1426-1435.
244. Kozel C*, et al.* (2016) Overexpression of eIF5 or its protein mimic 5MP perturbs eIF2 function and induces ATF4 translation through delayed re-initiation. *Nucleic Acids Res* 44(18):8704-8713.

245. Tang L*, et al.* (2017) Competition between translation initiation factor eIF5 and its mimic protein 5MP determines non-AUG initiation rate genome-wide. *Nucleic Acids Res* 45(20):11941-11953.
246. Benne R & Hershey JW (1976) Purification and characterization of initiation factor IF-E3 from rabbit reticulocytes. *Proc Natl Acad Sci U S A* 73(9):3005-3009.
247. Block KL, Vornlocher HP, & Hershey JW (1998) Characterization of cDNAs encoding the p44 and p35 subunits of human translation initiation factor eIF3. *J Biol Chem* 273(48):31901-31908.
248. Unbehaun A, Borukhov SI, Hellen CU, & Pestova TV (2004) Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. *Genes Dev* 18(24):3078-3093.
249. Cate JH (2017) Human eIF3: from 'blobology' to biological insight. *Philos Trans R Soc Lond B Biol Sci* 372(1716).
250. Hinnebusch AG (2006) eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* 31(10):553-562.
251. Asano K, Phan L, Anderson J, & Hinnebusch AG (1998) Complex formation by all five homologues of mammalian translation initiation factor 3 subunits from yeast Saccharomyces cerevisiae. *J Biol Chem* 273(29):18573-18585.
252. Browning KS*, et al.* (2001) Unified nomenclature for the subunits of eukaryotic initiation factor 3. *Trends Biochem Sci* 26(5):284.
253. Valasek L, Hasek J, Nielsen KH, & Hinnebusch AG (2001) Dual function of eIF3j/Hcr1p in processing 20 S pre-rRNA and translation initiation. *J Biol Chem* 276(46):43351-43360.
254. Valasek L, Phan L, Schoenfeld LW, Valaskova V, & Hinnebusch AG (2001) Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO J* 20(4):891-904.
255. Nielsen KH, Valasek L, Sykes C, Jivotovskaya A, & Hinnebusch AG (2006) Interaction of the RNP1 motif in PRT1 with HCR1 promotes 40S binding of eukaryotic initiation factor 3 in yeast. *Mol Cell Biol* 26(8):2984-2998.
256. Beznoskova P*, et al.* (2013) Translation initiation factors eIF3 and HCR1 control translation termination and stop codon read-through in yeast cells. *PLoS Genet* 9(11):e1003962.
257. Karaskova M*, et al.* (2012) Functional characterization of the role of the N-terminal domain of the c/Nip1 subunit of eukaryotic initiation factor 3 (eIF3) in AUG recognition. *J Biol Chem* 287(34):28420-28434.
258. Khoshnevis S*, et al.* (2014) Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes. *Nucleic Acids Res* 42(6):4123-4139.
259. Mohammad MP, Munzarova Pondelickova V, Zeman J, Gunisova S, & Valasek LS (2017) In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. *Nucleic Acids Res* 45(5):2658-2674.

260. Valasek L, *et al.* (2003) The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome in vivo. *Genes Dev* 17(6):786-799.
261. Pisarev AV, Kolupaeva VG, Yusupov MM, Hellen CU, & Pestova TV (2008) Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J* 27(11):1609-1621.
262. Kouba T, *et al.* (2012) Small ribosomal protein RPS0 stimulates translation initiation by mediating 40S-binding of eIF3 via its direct contact with the eIF3a/TIF32 subunit. *PLoS One* 7(7):e40464.
263. Kouba T, Rutkai E, Karaskova M, & Valasek L (2012) The eIF3c/NIP1 PCI domain interacts with RNA and RACK1/ASC1 and promotes assembly of translation preinitiation complexes. *Nucleic Acids Res* 40(6):2683-2699.
264. Hronova V, *et al.* (2017) Does eIF3 promote reinitiation after translation of short upstream ORFs also in mammalian cells? *RNA Biol* 14(12):1660-1667.
265. Pisarev AV, *et al.* (2006) Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev* 20(5):624-636.
266. Choi SK, Lee JH, Zoll WL, Merrick WC, & Dever TE (1998) Promotion of met-tRNAiMet binding to ribosomes by yIF2, a bacterial IF2 homolog in yeast. *Science* 280(5370):1757-1760.
267. Pestova TV, *et al.* (2000) The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403(6767):332-335.
268. Dever TE, *et al.* (2001) Universal translation initiation factor IF2/eIF5B. *Cold Spring Harb Symp Quant Biol* 66:417-424.
269. Roll-Mecak A, Cao C, Dever TE, & Burley SK (2000) X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. *Cell* 103(5):781-792.
270. Kuhle B & Ficner R (2014) eIF5B employs a novel domain release mechanism to catalyze ribosomal subunit joining. *EMBO J* 33(10):1177-1191.
271. Shin BS, *et al.* (2002) Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity. *Cell* 111(7):1015-1025.
272. Marintchev A, Kolupaeva VG, Pestova TV, & Wagner G (2003) Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: a new interaction between old partners. *Proc Natl Acad Sci U S A* 100(4):1535-1540.
273. Acker MG, Shin BS, Dever TE, & Lorsch JR (2006) Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. *J Biol Chem* 281(13):8469-8475.
274. Acker MG, *et al.* (2009) Kinetic analysis of late steps of eukaryotic translation initiation. *J Mol Biol* 385(2):491-506.
275. Zheng A, *et al.* (2014) X-ray structures of eIF5B and the eIF5B-eIF1A complex: the conformational flexibility of eIF5B is restricted on the ribosome by interaction with eIF1A. *Acta Crystallogr D Biol Crystallogr* 70(Pt 12):3090-3098.
276. Nag N, *et al.* (2016) eIF1A/eIF5B interaction network and its functions in translation initiation complex assembly and remodeling. *Nucleic Acids Res* 44(15):7441-7456.

277. Lee JH, *et al.* (2002) Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proc Natl Acad Sci U S A* 99(26):16689-16694.
278. Shin BS, *et al.* (2007) Intragenic suppressor mutations restore GTPase and translation functions of a eukaryotic initiation factor 5B switch II mutant. *Mol Cell Biol* 27(5):1677-1685.
279. Shin BS, *et al.* (2009) rRNA suppressor of a eukaryotic translation initiation factor 5B/initiation factor 2 mutant reveals a binding site for translational GTPases on the small ribosomal subunit. *Mol Cell Biol* 29(3):808-821.
280. Fringer JM, Acker MG, Fekete CA, Lorsch JR, & Dever TE (2007) Coupled release of eukaryotic translation initiation factors 5B and 1A from 80S ribosomes following subunit joining. *Mol Cell Biol* 27(6):2384-2397.
281. Cigan AM, Pabich EK, Feng L, & Donahue TF (1989) Yeast translation initiation suppressor sui2 encodes the alpha subunit of eukaryotic initiation factor 2 and shares sequence identity with the human alpha subunit. *Proc Natl Acad Sci U S A* 86(8):2784-2788.
282. Donahue TF, Cigan AM, Pabich EK, & Valavicius BC (1988) Mutations at a Zn(II) finger motif in the yeast eIF-2 beta gene alter ribosomal start-site selection during the scanning process. *Cell* 54(5):621-632.
283. Hannig EM, Cigan AM, Freeman BA, & Kinzy TG (1993) GCD11, a negative regulator of GCN4 expression, encodes the gamma subunit of eIF-2 in Saccharomyces cerevisiae. *Mol Cell Biol* 13(1):506-520.
284. Thompson GM, Pacheco E, Melo EO, & Castilho BA (2000) Conserved sequences in the beta subunit of archaeal and eukaryal translation initiation factor 2 (eIF2), absent from eIF5, mediate interaction with eIF2gamma. *Biochem J* 347 Pt 3:703-709.
285. Schmitt E, *et al.* (2012) Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat Struct Mol Biol* 19(4):450-454.
286. Yatime L, Mechulam Y, Blanquet S, & Schmitt E (2007) Structure of an archaeal heterotrimeric initiation factor 2 reveals a nucleotide state between the GTP and the GDP states. *Proc Natl Acad Sci U S A* 104(47):18445-18450.
287. Rajesh K, Iyer A, Suragani RN, & Ramaiah KV (2008) Intersubunit and interprotein interactions of alpha- and beta-subunits of human eIF2: Effect of phosphorylation. *Biochem Biophys Res Commun* 374(2):336-340.
288. Dhaliwal S & Hoffman DW (2003) The crystal structure of the N-terminal region of the alpha subunit of translation initiation factor 2 (eIF2alpha) from Saccharomyces cerevisiae provides a view of the loop containing serine 51, the target of the eIF2alpha-specific kinases. *J Mol Biol* 334(2):187-195.
289. Ito T, Marintchev A, & Wagner G (2004) Solution structure of human initiation factor eIF2alpha reveals homology to the elongation factor eEF1B. *Structure* 12(9):1693-1704.
290. Cho S & Hoffman DW (2002) Structure of the beta subunit of translation initiation factor 2 from the archaeon Methanococcus jannaschii: a representative of the eIF2beta/eIF5 family of proteins. *Biochemistry* 41(18):5730-5742.

291. Gutierrez P*, et al.* (2004) Structure of the archaeal translation initiation factor aIF2 beta from Methanobacterium thermoautotrophicum: implications for translation initiation. *Protein Sci* 13(3):659-667.
292. Sokabe M, Yao M, Sakai N, Toya S, & Tanaka I (2006) Structure of archaeal translational initiation factor 2 betagamma-GDP reveals significant conformational change of the beta-subunit and switch 1 region. *Proc Natl Acad Sci U S A* 103(35):13016-13021.
293. Schmitt E, Naveau M, & Mechulam Y (2010) Eukaryotic and archaeal translation initiation factor 2: a heterotrimeric tRNA carrier. *FEBS Lett* 584(2):405-412.
294. Schmitt E, Blanquet S, & Mechulam Y (2002) The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. *EMBO J* 21(7):1821-1832.
295. Hinnebusch AG (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* 59:407-450.
296. Sonenberg N & Hinnebusch AG (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136(4):731-745.
297. Asano K, Krishnamoorthy T, Phan L, Pavitt GD, & Hinnebusch AG (1999) Conserved bipartite motifs in yeast eIF5 and eIF2Bepsilon, GTPase-activating and GDP-GTP exchange factors in translation initiation, mediate binding to their common substrate eIF2. *EMBO J* 18(6):1673-1688.
298. Laurino JP, Thompson GM, Pacheco E, & Castilho BA (1999) The beta subunit of eukaryotic translation initiation factor 2 binds mRNA through the lysine repeats and a region comprising the C2-C2 motif. *Mol Cell Biol* 19(1):173-181.
299. Naveau M*, et al.* (2013) Roles of yeast eIF2alpha and eIF2beta subunits in the binding of the initiator methionyl-tRNA. *Nucleic Acids Res* 41(2):1047-1057.
300. Hashimoto NN, Carnevalli LS, & Castilho BA (2002) Translation initiation at non-AUG codons mediated by weakened association of eukaryotic initiation factor (eIF) 2 subunits. *Biochem J* 367(Pt 2):359-368.
301. Castilho-Valavicius B, Thompson GM, & Donahue TF (1992) Mutation analysis of the Cys-X2-Cys-X19-Cys-X2-Cys motif in the beta subunit of eukaryotic translation initiation factor 2. *Gene Expr* 2(3):297-309.
302. Thakur A, Marler L, & Hinnebusch AG (2019) A network of eIF2beta interactions with eIF1 and Met-tRNAi promotes accurate start codon selection by the translation preinitiation complex. *Nucleic Acids Res* 47(5):2574-2593.
303. Chen SJ, Lin G, Chang KJ, Yeh LS, & Wang CC (2008) Translational efficiency of a non-AUG initiation codon is significantly affected by its sequence context in yeast. *J Biol Chem* 283(6):3173-3180.
304. Kozak M (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res* 12(2):857-872.
305. Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15(20):8125-8148.
306. Kozak M (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* 266(30):19867-19870.
307. Kozak M (1984) Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo. *Nature* 308(5956):241-246.

308. Joshi CP, Zhou H, Huang X, & Chiang VL (1997) Context sequences of translation initiation codon in plants. *Plant Mol Biol* 35(6):993-1001.
309. Hamilton R, Watanabe CK, & de Boer HA (1987) Compilation and comparison of the sequence context around the AUG startcodons in Saccharomyces cerevisiae mRNAs. *Nucleic Acids Res* 15(8):3581-3593.
310. Dvir S*, et al.* (2013) Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc Natl Acad Sci U S A* 110(30):E2792-2801.
311. Li J, Liang Q, Song W, & Marchisio MA (2017) Nucleotides upstream of the Kozak sequence strongly influence gene expression in the yeast S. cerevisiae. *J Biol Eng* 11:25.
312. Lee S*, et al.* (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A* 109(37):E2424-2432.
313. Acevedo JM, Hoermann B, Schlimbach T, & Teleman AA (2018) Changes in global translation elongation or initiation rates shape the proteome via the Kozak sequence. *Sci Rep* 8(1):4018.
314. Visweswaraiah J, Pittman Y, Dever TE, & Hinnebusch AG (2015) The beta-hairpin of 40S exit channel protein Rps5/uS7 promotes efficient and accurate translation initiation in vivo. *Elife* 4:e07939.
315. Sharifulin D*, et al.* (2013) Ribosomal protein S5e is implicated in translation initiation through its interaction with the N-terminal domain of initiation factor eIF2alpha. *Chembiochem* 14(16):2136-2143.
316. Visweswaraiah J & Hinnebusch AG (2017) Interface between 40S exit channel protein uS7/Rps5 and eIF2alpha modulates start codon recognition in vivo. *Elife* 6:e22572.
317. Sharifulin D*, et al.* (2012) A central fragment of ribosomal protein S26 containing the eukaryote-specific motif YxxPKxYxK is a key component of the ribosomal binding site of mRNA region 5' of the E site codon. *Nucleic Acids Res* 40(7):3056-3065.
318. Ferretti MB, Ghalei H, Ward EA, Potts EL, & Karbstein K (2017) Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements. *Nat Struct Mol Biol* 24(9):700-707.
319. Hinnebusch AG & Lorsch JR (2012) The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* 4(10):a011544.
320. Jackson RJ, Hellen CU, & Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11(2):113-127.
321. Voigts-Hoffmann F, Klinge S, & Ban N (2012) Structural insights into eukaryotic ribosomes and the initiation of translation. *Curr Opin Struct Biol* 22(6):768-777.
322. Pestova TV, Borukhov SI, & Hellen CU (1998) Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 394(6696):854-859.
323. Valasek LS (2012) 'Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr Protein Pept Sci* 13(4):305-330.

324. Erzberger JP*, et al.* (2014) Molecular architecture of the 40SeIF1eIF3 translation initiation complex. *Cell* 158(5):1123-1135.
325. Siridechadilok B, Fraser CS, Hall RJ, Doudna JA, & Nogales E (2005) Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* 310(5753):1513-1515.
326. Scheres SH (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180(3):519-530.
327. Mindell JA & Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334-347.
328. Bai XC, Fernandez IS, McMullan G, & Scheres SH (2013) Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* 2:e00461.
329. Scheres SH & Chen S (2012) Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9(9):853-854.
330. Kucukelbir A, Sigworth FJ, & Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods* 11(1):63-65.
331. Rosenthal PB & Henderson R (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* 333(4):721-745.
332. Pettersen EF*, et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
333. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
334. Brown A*, et al.* (2015) Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr D Biol Crystallogr* 71(Pt 1):136-153.
335. Stolboushkina E*, et al.* (2008) Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. *J Mol Biol* 382(3):680-691.
336. Cuchalova L*, et al.* (2010) The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. *Mol Cell Biol* 30(19):4671-4686.
337. Herrmannova A*, et al.* (2012) Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. *Nucleic Acids Res* 40(5):2294-2311.
338. Reibarkh M*, et al.* (2008) Eukaryotic initiation factor (eIF) 1 carries two distinct eIF5-binding faces important for multifactor assembly and AUG selection. *J Biol Chem* 283(2):1094-1103.
339. Dong Z, Qi J, Peng H, Liu J, & Zhang JT (2013) Spectrin domain of eukaryotic initiation factor 3a is the docking site for formation of the a:b:i:g subcomplex. *J Biol Chem* 288(39):27951-27959.
340. Hinnebusch AG (2017) Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci* 42(8):589-611.
341. Donahue T (2000) Genetic approaches to translation initiation in *Saccharomyces cerevisiae*. *Translational Control of Gene Expression*, eds Sonenberg N, Hershey

JWB, & Mathews MB (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp 487-502.
342. Martin-Marcos P*, et al.* (2013) beta-hairpin loop of eIF1 mediates 40S ribosome binding to regulate initiator tRNAMet recruitment and accuracy of AUG selection in vivo. *J Biol Chem* 288(38):27546-62.
343. Cross FR (1997) 'Marker swap' plasmids: convenient tools for budding yeast molecular genetics. *Yeast* 13(7):647-653.
344. Moehle CM & Hinnebusch AG (1991) Association of RAP1 binding sites with stringent control of ribosomal protein gene transcription in *Saccharomyces cerevisiae*. *Mol Cell Biol* 11:2723-2735.
345. Reid GA & Schatz G (1982) Import of proteins into mitochondria. Yeast cells grown in the presence of carbonyl cyanide m-chlorophenylhydrazone accumulate massive amounts of some mitochondrial precursor polypeptides. *J Biol Chem* 257(21):13056-13061.
346. Olsen DS*, et al.* (2003) Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment *in vivo*. *EMBO J* 22:193-204.
347. Acker MG, Kolitz SE, Mitchell SF, Nanda JS, & Lorsch JR (2007) Reconstitution of yeast translation initiation. *Methods Enzymol* 430:111-145.
348. Grant CM, Miller PF, & Hinnebusch AG (1994) Requirements for intercistronic distance and level of eIF-2 activity in reinitiation on GCN4 mRNA varies with the downstream cistron. *Mol Cell Biol* 14:2616-2628.
349. Huang H, Yoon H, Hannig EM, & Donahue TF (1997) GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *Genes Dev* 11:2396-2413.
350. Alone PV, Cao C, & Dever TE (2008) Translation initiation factor 2gamma mutant alters start codon selection independent of Met-tRNA binding. *Mol Cell Biol* 28(22):6877-6888.
351. De Angioletti M, Lacerra G, Sabato V, & Carestia C (2004) Beta+45 G --> C: a novel silent beta-thalassaemia mutation, the first in the Kozak sequence. *Br J Haematol* 124(2):224-231.
352. Yusupova GZ, Yusupov MM, Cate JH, & Noller HF (2001) The path of messenger RNA through the ribosome. *Cell* 106(2):233-241.
353. Graifer D*, et al.* (2004) Variable and conserved elements of human ribosomes surrounding the mRNA at the decoding and upstream sites. *Nucleic Acids Res* 32(11):3282-3293.
354. Longtine MS*, et al.* (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* 14(10):953-961.
355. Sikorski RS & Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* 122(1):19-27.
356. Moehle CM & Hinnebusch AG (1991) Association of RAP1 binding sites with stringent control of ribosomal protein gene transcription in Saccharomyces cerevisiae. *Mol Cell Biol* 11(5):2723-2735.
357. Fassio CA, Schofield BJ, Seiser RM, Johnson AW, & Lycan DE (2010) Dominant mutations in the late 40S biogenesis factor Ltv1 affect cytoplasmic

maturation of the small ribosomal subunit in Saccharomyces cerevisiae. *Genetics* 185(1):199-209.
358. Dorris DR, Erickson FL, & Hannig EM (1995) Mutations in GCD11, the structural gene for eIF-2 gamma in yeast, alter translational regulation of GCN4 and the selection of the start site for protein synthesis. *EMBO J* 14(10):2239-2249.
359. Christianson TW, Sikorski RS, Dante M, Shero JH, & Hieter P (1992) Multifunctional yeast high-copy-number shuttle vectors. *Gene* 110(1):119-122.
360. Donahue TF & Cigan AM (1988) Genetic selection for mutations that reduce or abolish ribosomal recognition of the HIS4 translational initiator region. *Mol Cell Biol* 8(7):2955-2963.
361. Gietz RD & Sugino A (1988) New yeast-Escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* 74(2):527-534.
362. Parent SA, Fenimore CM, & Bostian KA (1985) Vector systems for the expression, analysis and cloning of DNA sequences in S. cerevisiae. *Yeast* 1(2):83-138.
363. Hinnebusch AG (1985) A hierarchy of trans-acting factors modulates translation of an activator of amino acid biosynthetic genes in Saccharomyces cerevisiae. *Mol Cell Biol* 5(9):2349-2360.
364. Dever TE, Yang W, Astrom S, Bystrom AS, & Hinnebusch AG (1995) Modulation of tRNA(iMet), eIF-2, and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of eIF-2.GTP.Met-tRNA(iMet) ternary complexes. *Mol Cell Biol* 15(11):6351-6363.
365. Pavitt GD, Ramaiah KV, Kimball SR, & Hinnebusch AG (1998) eIF2 independently binds two distinct eIF2B subcomplexes that catalyze and regulate guanine-nucleotide exchange. *Genes Dev* 12(4):514-526.
366. Gomez E, Mohammad SS, & Pavitt GD (2002) Characterization of the minimal catalytic domain within eIF2B: the guanine-nucleotide exchange factor for translation initiation. *EMBO J* 21(19):5292-5301.

# 9. CV

Laura Marler
marler.laura1@gmail.com

## Education

**Johns Hopkins University, National Institutes of Health Graduate Partnership Program,** Bethesda, MD
Graduated 2019, PhD in Cell, Molecular, and Developmental Biology
**Furman University**, Greenville, South Carolina
Graduated 2005, B.S. in chemistry with an emphasis in biochemistry, *cum laude*
**Tennessee High School**, Bristol, Tennessee
Graduated 2001, valedictorian

## Research Experience

2013 – 2019, **Section on Nutrient Control of Gene Expression, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health,** Bethesda, Maryland, Alan Hinnebusch lab
Graduate Student, Role of ribosomal proteins and initiation factor eIF2β in accurate recognition of translation initiation start sites

Spring 2013, **Department of Biology, Johns Hopkins University,** Baltimore, Maryland, Kyle Cunningham lab
Graduate Student rotation, Role of calcineurin in cell death

Fall 2012, **Department of Embryology, Carnegie Institute of Washington**, Baltimore, Maryland, Nicholas Ingolia lab
Graduate Student rotation, Development of a dual luciferase assay to probe alternate translation initiation sites

Fall 2012, **Zanvyll-Kreiger Mind Brain Institute, Johns Hopkins University**, Baltimore, Maryland, Hey-Kyoung Lee lab
Graduate Student rotation, Synaptic plasticity in response to dark exposure

Summer 2012, **Laboratory of Molecular Pharmacology, National Cancer Institute**, Bethesda, Maryland, Yves Pommier lab
Graduate Student rotation, Inhibitors of HIV integrase

2007-2012**, College of Pharmacy, University of Hawaii at Hilo,** Hilo, Hawaii, John Pezzuto lab
Research Technician, Natural Inhibitors of Carcinogenesis
Bioassay-guided fractionation and mechanistic evaluation of compounds demonstrating aromatase inhibition and induction of quinone reductase 1

2005-2006, **National Flow Cytometry Resource, Los Alamos National Laboratory**, Los Alamos, New Mexico, Steve Graves lab
Post-baccalaureate Fellow, Kinetics of protease cleavage in *B. anthracis* and applications to other toxins

2004-2005, **Furman University**, Greenville, South Carolina, Moses Lee lab
Student Researcher, Biophysical analysis of polyamide minor groove binders

Summer 2004, **Influenza, Molecular Genetics Section, Centers for Disease Control**, Atlanta, Georgia, Ruben Donis lab
Student Researcher, Sequence analysis of heamagluttination inhibition assay results

Summer 2003, **Physical Chemistry and Spectroscopy, Los Alamos National Laboratory**, Los Alamos, New Mexico, Jeanne Robinson lab
Student Researcher, Second harmonic microscopy of thin films

## Teaching Experience:

FAES Instructor, BIOL262, "Research Tools for Studying Disease", Fall, 2016, 2017

FAES Biotechnology Workshop Instructor, 2015

Summer Intern Journal Club leader, Office of Intramural Training and Education, National Institutes of Health, 2014 – 2017 (led six week courses of 10 – 15 students)

Cell Biology Lab teaching assistant, Biology Department, Johns Hopkins University, Spring 2013, Spring 2014 (led one lab section of 20 – 25 students for two sixteen week terms)

One-on-one online tutor, biology and chemistry courses, grades 7 – 12, tutor.com, 2010 – 2012.

Chemistry tutor, group tutor for 10-12 students, organic chemistry and biochemistry, Chemistry Department, University of Hawaii at Hilo, 2009 – 2010.

Chemistry teaching assistant, general chemistry, spectroscopy, and organic chemistry labs, 2-3 lab sections of 20 – 30 students per year, Chemistry Department, Furman University, 2003 – 2005.

## Scientific Outreach:

NIH Graduate Student Symposium Planning Committee member, National Institutes of Health, 2017 – 2018.

Mentor, Bnos Yisroel Bridge Program, National Institutes of Health, Summer, 2017.

Postbaccalaureate Poster Day Judge, National Institutes of Health, 2016 - 2018.

S.T.E.M. Fair judge, Gwynn Park High School, Brandywine, Maryland, 2015 - 2016.

National Gay Blood Drive local organizer, Washington, D.C., 2013 – 2015.

MInDS (Mentoring to Inspire Diversity in Science) mentor, Baltimore, Maryland, 2012 – 2013.

## Additional Training:

Regulatory Affairs and FDA Regulation, 15 week course, Foundation for Advanced Education in the Sciences, Spring 2018.

Translational Science Training Program, 2 day intensive training course on the drug discovery and development process, including FDA approvals, Office of Intramural Training and Education, NIH, Spring 2018.

Business of Science, a 4 part certification course on communication, building effective teams, financial literacy, and project management, National Institute of Child Health and Human Development, NIH, Spring 2018.

Management Bootcamp, an intense 2 day management certification, Office of Intramural Training and Education, NIH, Spring 2018.

Workplace Dynamics, a 5 part course on self-awareness, conflict, feedback, team skills, and diversity in the workplace, Office of Intramural Training and Education, NIH, 2017.

Scientists Teaching Science, semester-long course, Office of Intramural Training and Education, NIH, 2015.

## Publications:

**Marler L***, Thakur A*, Nanda JS, Lorsch JR, Hinnebusch AG. Interactions at the eIF1/eIF2β interface contribute to start codon selection accuracy in eukaryotic translation. *Nucleic Acids Research* **2019**, 47(5):2574-2593. *Authors contributed equally.

Metifiot M, Johnson B, **Marler L**, Zhao XZ, Burke T, Marchand C, Hughes SH, Pommier Y. Selectivity for 3'-processing over strand-transfer and susceptibility to clinical resistance of HIV-1 integrase inhibitors are both driven by key enzyme-DNA interactions in the active site. *Nucleic Acids Research* **2016**, 44(14):6896-6906.

Fatima N, Kondratyuk TP, Park EJ, **Marler L**, Jadoon M, Qazi MA, Mehboob Mirza H, Khan I, Atiq N, Chang LC, Ahmed S, Pezzuto JM. Endophytic fungi associated with Taxus fuana (West Himalayan Yew) of Pakistan: potential bio-resources for cancer chemopreventive agents. *Pharmaceutical Biology* **2016**, 54(11):2547-2554.

Llacer JL, Hussain T, **Marler L**, Aitken CE, Thakur A, Lorsch JR, Hinnebusch AG, Ramakrishnan V. Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Molecular Cell* **2015**, 59(3):399-412.

Garcia ME, Nicotra VE, Oberti JC, Rios-Luci C, Leon LG, **Marler L**, Li G, Pezzuto JM, van Breemen RB, Padron JM, Hueso-Falcon I, Estevez-Braun A. Antiproliferative and quinone reductase-inducing activities of withanolides derivatives. *European Journal of Medicinal Chemistry* **2014**, 82:68–81.

Carlson S, **Marler L**, Nam S-J, Santarsiero BD, Pezzuto JM, Murphy BT. Potential chemopreventive activity of a new macrolide antibiotic from a marine-derived *Micromonospora* sp. *Marine Drugs* **2013**, 11(4):1152-1161.

**Marler L** and Pezzuto JM. Nutritional Phytochemicals and the Management of Chronic Inflammation. *In:* T. Kong (ed.), *Inflammation, Oxidative Stress, and Cancer: Dietary Approaches for Cancer Prevention*, Taylor and Francis Books, **2013**.

Haq IU, Mizra B, Kondratyuk TP, Park E-J, Burns BE, **Marler L**, Pezzuto JM. Preliminary evaluation for cancer chemopreventive and cytotoxic potential of naturally growing ethnobotanically selected plants of Pakistan. *Pharmaceutical Biology* **2013**, 51(3):316-328.

Mayhoub AS, **Marler L**, Kondratyuk TP, Park E-J, Pezzuto JM, Cushman M. Optimization of Thiazole Analogues of Resveratrol for induction of NADP(H):Quinone Reductase 1 (QR1). *Bioorganic and Medicinal Chemistry* **2012**, 20(24); 7030-7039.

Luqman S, Meena A, Singh P, Kondratyuk TP, **Marler L**, Pezzuto JM, Negi AS. Neoflavonoids and tetrahydroquinolones as possible cancer chemopreventive agents. *Chemical and Biological Drug Design* **2012**, 80(4):616-624.

Mayhoub AS, **Marler L**, Kondratyuk T, Park E-J, Pezzuto JM, Cushman M. Optimization of the Aromatase Inhibitory Activities of Pyridylthiazole Analogues of Resveratrol. *Bioorganic and Medicinal Chemistry* **2012**, 20(7):2427-2434.

Mayhoub AS, **Marler L**, Kondratyuk T, Park E-J, Pezzuto JM, Cushman M. Optimizing Thidiazole Analogues of Resveratrol versus Three Chemopreventive Targets. *Bioorganic and Medicinal Chemistry* **2012,** 20(1):510-520.

Shen L, Park E-J, Kondratyuk TP, Guendisch D, **Marler L**, Pezzuto JM, Wright AD, Sun D. Synthesis of callophycin A analogues and evaluation as potential chemopreventive and anticancer agents. *Bioorganic and Medicinal Chemistry* **2011,** 19(21):6182-6195.

Luqman S, Meena A, **Marler L**, Kondratyuk TP, Pezzuto JM. Suppression of tumor necrosis factor-alpha induced NFkB Activation and Aromatase Activity by Capsaicin and its Analogue Capsazepine. *Journal of Medicinal Food* **2011,** 14(11):1344-1351.

Kondratyuk TP, Park E-J, **Marler L**, Ahn S, Yuan Y, Choi Y, van Breemen RB, Sun B, Hoshino J, Cushman M, Jermihov KC, Mesecar AD, Grubbs CJ, Pezzuto JM. Resveratrol derivatives as promising chemopreventive agents with improved potency and selectivity. *Molecular Nutrition and Food Research* **2011,** Aug;55(8):1249-1265.

Yang J, Kondratyuk TP, Jermihov KC, **Marler L**, Qiu X, Choi Y, Cao H, Yu R, Sturdy M, Huang R, Liu Y, Wang L, Mesecar AD, van Breemen RB, Pezzuto JM, Fong HHS, Chen Y, Zhang H. Bioactive compounds from the fern *Lepisorus contortus*. *Journal of Natural Products* **2011,** 74(2):129-136.

**Marler L**, Maiti A, Conda-Sheridan M, Cushman M, Chen L, Huang K, van Breemen RB, Grubbs C, Pezzuto JM. Cancer Chemopreventive Potential of Aromathecins and Phenazines, Novel Natural Product Derivatives. *Anticancer Research* **2010,** 30(12):4873-4882.

Conda-Sheridan M, **Marler L**, Park E-J, Kondratyuk TP, Jermihov K, Mesecar AD, Pezzuto JM, Fenical W, Cushman M. Potential chemopreventive agents based on the structure of the lead compound 2-bromo-1-hydroxyphenazine, isolated from Streptomyces species, strain CNS284. *Journal of Medicinal Chemistry* **2010,** 53(24):8688-8699.

Hoshino J, Park EJ, Kondratyuk TP, **Marler L**, Pezzuto JM, Cushman M. Selective Synthesis and Biological Evaluation of Sulfate-Conjugated Resveratrol Metabolites. *Journal of Medicinal Chemistry* **2010**, 53(13):5033-5043.

Sun B, Hoshino J, Jermihov K, **Marler L**, Pezzuto JM, Mesecar AD, Cushman M. Design, Synthesis, and Biological Evaluation of Resveratrol Derivatives as Aromatase and Quinone Reductase 2 Inhibitors for Chemoprevention of Cancer. *Bioorganic and Medicinal Chemistry* **2010**, 18(14):5352-5366.

Nam SJ, Guadencio SP, Kauffman CA, Jensen PR, Kondratyuk TP, **Marler L**, Pezzuto JM, Fenical W. Fijiolides A and B, Inhibitors of TNF-α Induced NFκB Activation, from a Marine-derived Sediment Bacterium of Genus *Nocardiopsis*. *Journal of Natural Products* **2010**, 73(6):1080-1086.

Yang JH, Kondratyuk TP, **Marler L**, Qiu X, Choi Y, Cao H, Yu R, Sturdy M, Pegan S, Liu Y, Wang LQ, Mesecar AD, van Breemen RB, Pezzuto JM, Fong HHS, Chen YG, Zhang HJ. Isolation and Evaluation of Kaempferol Glycosides from the Fern *Neocheiropteris palmatopedata*. *Phytochemistry* **2010,** 71(5-6):641-647.

Cuendet M, Guo J, Luo Y, Chen S, Oteham CP, Moon RC, van Breemen RB, **Marler L**, Pezzuto JM. Cancer Chemopreventive Activity and Metabolism of Isoliquiritigenin, a Compound Found in Licorice. *Cancer Prevention Research* **2010**, 3(2):221-232.

Schupp P, Kohlert-Schupp C, Whitefield S, Engemann A, Rohde S, Hemscheidt T, Pezzuto JM, Kondratyuk TP, Park EJ, **Marler L**, Rostama B, Wright AD. Cancer Chemopreventive and Anti-Cancer Evaluation of Extracts and Fractions from Marine Macro- and Micro-organisms Collected from Twilight Zone Waters Around Guam. *Natural Product Communications* **2009,** 4(12):1717-1728.

Ghufran MA, Qureshi RA, Batool A, Kondratyuk TP, Guilford JM, **Marler L**, Chang LC, Pezzuto JM. Evaluation of Selected Indigenous Medicinal Plants from the Western Himalayas for Cytotoxicity and as Potential Cancer Chemopreventive Agents. *Pharmaceutical Biology* **2009**, 47(6): 533-538.

Maiti A, Reddy PVN, Sturdy M, **Marler L**, Pegan SD, Mesecar AD, Pezzuto JM, Cushman M. Synthesis of Casimiroin and Optimization of its Quinone Reductase 2 and Aromatase Inhibitory Activities. *Journal of Medicinal Chemistry* **2009**, 52(7): 1873-1884.

## Abstracts Presented:

New role for translation initiation factor eIF2β in promoting PIC assembly and accurate start codon selection *in vivo*. **Laura Marler**, Alan Hinnebusch. *Translational Control,* **2018**, Cold Spring Harbor, New York.

eIF1/Met-tRNA$_i$ and eIF1/eIF2β interactions in the open preinitiation complex and eIF2α/mRNA contacts in the closed state exert opposing effects on translation initiation accuracy. Anil Thakur, **Laura Marler**, Alan Hinnebusch. *EMBL Protein Synthesis and Translational Control*, **2017**, Heidelberg, Germany.

40S ribosomal proteins US7/RPS5 and RPS26E in the mRNA exit channel play critical roles in accurate start codon selection during translation initiation *in vivo*. **Laura Marler**, Jyothsna Visweswaraiah, Alan Hinnebusch. *Translational Control*, **2016**, Cold Spring Harbor, New York.

New role for translation initiation factor eIF2β in promoting ribosomal scanning and accurate start codon selection *in vivo*. **Laura Marler**, Anil Thakur, Jose L. Llacer, Tanweer Hussain, Colin Aitken, Jon Lorsch, Venki Ramakrishnan, Alan Hinnebusch. *EMBL Protein Synthesis and Translational Control*, **2015**, Heidelberg, Germany.

Thiazole and thiadiazole derivatives of resveratrol as inducers of quinone reductase 1. **Laura Marler**, Abdelrahman S. Mayhoub, Mark Cushman, John M. Pezzuto. *50$^{th}$ Anniversary Meeting of the Phytochemical Society of North America*, **2011**, Kona, Hawaii.

2,4-Diaryl-1,3-thiazoles: Novel aromatase inhibitors. Abdelrahman S. Mayhoub, **Laura Marler**, Tamara Kondratyuk, Eun-Jung Park, John M. Pezzuto, Mark Cushman. *Third Annual Breast Cancer Discovery Group Retreat,* **2011**, Purdue University.

Cancer Chemopreventive Potential of Marine-derived Phenazines. **Laura Marler**, John M. Pezzuto, Ratnakar N. Asolkar, William Fenical, Martin Conda-Sheridan, Mark Cushman. *Pacifichem*, **2010**, Honolulu, Hawaii.

Specificity of Casimiroin Analogs as Inhibitors of Aromatase. **Laura Marler**, Arup Maiti, Tamara Kondratyuk, Mark Cushman, John M. Pezzuto. *American Society of Pharmacognosy*, **2010**, St. Petersburg, Florida.

Marine Phenazines Inhibit Invasiveness of MDA-MB-231 Breast Cancer Cells through Extracellular Matrix. Tamara P. Kondratyuk, Eun-Jung Park, **Laura Marler**, Brittany E. Burns, Ratnakar N. Asolkar, William Fenical, John M. Pezzuto. *American Society of Pharmacognosy*, **2010**, St. Petersburg, Florida.

Elements of Structure Elucidation of Isolates from the Twilight Zone Sponge *SUBEREA* SP. Collected from Waters around Guam. Anthony D. Wright, Peter Schupp, Claudia Kohlert-Schupp, John M. Pezzuto, Tamara P. Kondratyuk, Eun-Jung Park, **Laura Marler**. *American Society of Pharmacognosy*, **2010**, St. Petersburg, Florida.

Cancer Chemopreventive Potential of Selected Medicinal Plants from Pakistan. Ihsan ul Haq, Bushra Mirza, Eun-Jung Park, Brittany E. Burns, **Laura Marler**, Tamara P. Kondratyuk, John M. Pezzuto. *American Society of Pharmacognosy*, **2010**, St. Petersburg, Florida.

Aromatase Inhibition and Chemopreventive Potential of Novel Resveratrol Derivatives. **Laura Marler**, Mark Cushman, Sun Bin, Andrew D. Mesecar, Katie Jermihov, Richard C. Van Breemen, John M. Pezzuto. *AACR International Conference on Frontiers in Cancer Prevention Research*, **2009**, Houston, Texas.

Chemopreventive Potential of Natural Products Isolated from Alchornea glandulosa, Pterogyne nitens and its Semisyntheitc Analogs. Mauro C. C. de Morais, **Laura Marler**, Suaib Luqman, Tamara P. Kondratyuk, Maicon S. Petronio, Luis O. Regasini, Dulce H. S. Silva, Vanderlan S. Bolzani, Christiane P. Soares, John M. Pezzuto. *7th International Congress of Pharmaceutical Sciences (CIFARP)*, **2009**, Riberao Preto, Brazil.

Inhibition of NFkB activation and aromatase activity by vanilloids: an in vitro and in silico study. Suaib Luqman, Abha Meena, **Laura Marler**, Tamara P. Kondratyuk, John M. Pezzuto. *Eighth International Congress on Targeted Therapies in Cancer*, **2009**, Washington, D.C.

Cancer Chemopreventive Potential of Novel Natural Product Derivatives. **Laura Marler**, Beau Rostama, Mark Cushman, Richard van Breemen, Clinton Grubbs, John M. Pezzuto. *50$^{th}$ Anniversary Meeting of the American Society of Pharmacognosy*, **2009**, Honolulu, Hawaii.

Mechanism Based Screening of Resveratrol Derivatives as Potential Cancer Chemopreventive Agents. Tamara P. Kondratyuk, Eun-Jung Park, **Laura Marler**,

Bahman Rostama, John M. Pezzuto, Juma Hoshino, Mark Cushman. *50th Anniversary Meeting of the American Society of Pharmacognosy*, **2009**, Honolulu, Hawaii.

Chemoprotective and Anticancer Screening of Extracts from Marine Macro- and Micro-organisms from Waters around Guam. Peter Schupp, Claudia Kohlert-Schupp, John M. Pezzuto, Tamara P. Kondratyuk, Eun-Jung Park, **Laura Marler**, Anthony D. Wright. *50th Anniversary Meeting of the American Society of Pharmacognosy*, **2009**, Honolulu, Hawaii.

Evaluation of Indigenous Medicinal Plants from the Western Himalayas as Potential Cancer Chemopreventive Agents. Ghufran MA, Qureshi RA, Batool A, Kondratyuk TP, Guilford J, **Marler L**, Chang LC, Pezzuto JM. *Proceedings of the 31st Annual Conference of the Society of Ethnobiology*, **2008**, Fayetteville, Arkansas.

The Marine Environment as a Resource for the Discovery of Chemopreventive Agents. Tamara P. Kondratyuk, **Laura Marler**, Jackie Guilford, Katherine Maloney, William Fenical, John M. Pezzuto. *89th Annual Meeting of the AAAS Pacific Division*, **2008**, Waimea, Hawaii.

Biophysical Studies of Pyrrole/Imidazole Tetraamides Utilizing an Ethidium Bromide Displacement Assay. Andrea Bryant, Kellie King, **Laura Marler**, Lindsay Stollings, Karen Buchmueller, Hilary Mackay, Toni Brown, Moses Lee. *South-Eastern Regional Meeting of the American Chemical Society*, **2005**, Atlanta, Georgia.